UNDERSTANDING FACTORS INFLUENCING NON-INDUSTRIAL PRIVATE FOREST OWNERS' DECISIONS IN PLANTING AND HARVESTING TREES: A CASE STUDY IN THAI NGUYEN PROVINCE, VIETNAM

A thesis submitted in partial fulfilment of the requirements for

the Degree of Doctor of Philosophy in Forestry

by Viet Anh Dong

University of Canterbury

2018

Table of Contents

LIST OF TABLES

LIST OF FIGURES

LIST OF EQUATIONS

LIST OF ACRONYMS

AIC: Akaike Information Criterion.

FLA: Forest Land Allocation.

HEC: Human Ethics Committee of the University of Canterbury.

IAD: Institutional Analysis Development

LASSO: Least Absolute Skrinkage and Selection Operator.

NIPF owners: Non-Industrial Private Forest Owners.

PFES: Payment for Forest Environmental Services.

REDD+: Reducing Emissions from Deforestation and forest Degradation.

SFE: State Forestry Enterprises.

ACKNOWLEDGEMENT

I would firstly like to express my sincere gratitude to my primary supervisor, Professor Bruce Manley, for his knowledge, lessons, advice and support from the very early stages of my PhD tenure at School of Forestry. I also would like to thank to my second supervisor, Associate Professor Luis Apiolaza, who patiently taught me statistics and data analysis skills. My analysis skills have been significantly improved after years of working with you. I was very lucky to have worked with both of you and I will do my very best to translate what I had learnt from you into practice.

I would like to thank the forest owners who voluntarily participated in this research. This research would not be possible without their support and valuable information. I also would like to extend my thank to my colleagues at the Thai Nguyen Provincial Department of Agriculture and Rural Development, and my research assistants, Ms Ha, Ms Châu and Mr Sơn, who helped me to carry out the surveys and collect the data.

I would like to acknowledge the New Zealand Government for funding my studies in the form of a scholarship (New Zealand ASEAN Scholars scholarship), the School of Forestry (Postgraduate Scholarship 2017) for funding my surveys, and the Vietnamese Government for allowing me to take study leave in order to pursue my PhD for the last three years.

At the School of Forestry, I would like to thank all academic and staff members for treating me well during my study. I am particularly grateful for the support of Jeanette Allen and David Clark, who bent over backwards to provide me with assistance whenever I needed it.

Many thanks to Katinia Makaafi (student care team member at University of Canterbury), my friends, especially Sanjay Sekaran (Singapore) and Dung Duong Tien (Vietnam), who have always aided me punctually within the required timeframes that I had set. My friends at Ilam Apartments flat 18 and the research fellows at School of Forestry provided a very supportive and fun environment for my work. It was great to share these years with you.

And most of all, a very special thanks to my wife and family who always unconditionally loved and believed in me. Without you, none of this would have been possible and I am blessed to have you in my life.

Christchurch, Canterbury, New Zealand

20th November 2018

ABSTRACT

In the last few decades, the forest management system in Vietnam has switched from being centrally-planned and has become market-oriented. Under the current management model, private ownership is recognised and forests and forestland are allocated to non-industrial private forest owners, who are one of the biggest groups of forest owners in the country. After the transformation of the management model, the forest cover of the country shifted from net deforestation to net reforestation. However, the contribution of forestland allocation to this net reforestation is still a topic of debate.

The central argument for developing this research is that the most important aspect of forest management is the understanding of forest owners' behaviour in planting and harvesting trees. This purpose of this research is therefore to improve the understanding of non-industrial private forest owners' decisions in planting and harvesting trees by performing a case study in Thai Nguyen province, Vietnam.

Institutional Analysis Development (IAD) was used as the principal framework to guide the direction of the research. The data for the study was collected through two interview surveys in 2015 and 2017. Logistic and linear mixed-effects models were used to identify factors affecting the forest owners' decisions in planting and harvesting trees. The best subset selection, Ridge regression and Least Absolute Shrinkage and Selection Operator (LASSO) techniques were used to obtain multiple-predictor models and to quantify the importance of individual predictors in the models.

The results of this study indicate that the most important factor affecting the afforestation intensity of the forest owners is the forest owners' perception of their forestland. It was found that the forest owners who owned forestland in order to generate cash income from forestry activities were more likely to plant trees. Conversely, it was found that the forest owners who considered forestland as an investment tended not to plant trees. Additionally, the total number of forestland plots, annual income and awareness of the government subsidy for establishing forests were positively correlated with afforestation intensity. Meanwhile, age, level of education, total cropland area of the forest owners and total number of people in the workforce in the forest owners' family were negatively correlated with afforestation intensity.

The harvest intensity model was developed with respect to *Acacia mangium* because this is the dominant species planted in the province. The best-fit model indicates that tree age and

gender of the forest owners were correlated with harvest intensity. Timber price was positively correlated with harvest intensity. The cost of harvesting also plays a role in describing harvest intensity.

The results of this study suggest that the factors affecting the decisions of the forest owners are diverse. Therefore, it is necessary to develop management tools that allows forestry policy makers to (i) understand their policy-targeted audiences, (ii) test the impact of their policies during the policy-design stage, and (iii) receive feedback from their targeted audiences by observing changes in society. The approach developed in this comprehensive study is applicable for this purpose and is easily generalisable to a wide range of regions or countries.

CHAPTER 1: INTRODUCTION

Forestry governance and sustainability are two of the most important topics of debate in the realm of forestry, at both the global and national level. In the last few decades, many forms of forestry governance have been introduced at various levels of government, all the way from the multinational to the local level. The governmental institutions employ various management regimes in forest management including decentralised management, market-oriented management and citizen participation, with the ultimate goal of sustainably managing forests and forestland. This is often combined with efforts in combating the climate change phenomenon by finding a common understanding between the governments and their target populace.

In the last few decades, the forestry sector of Vietnam has displayed a pattern of management transformation. The country shifted from net deforestation to net reforestation level (Dang et al. 2012; Lambini et al. 2018), which is remarkable considering the fact that Vietnam had lost about half of its total forested area between 1945 and the early 1990s. One of the reasons leading to the net deforestation was the politico-economic model of the country at the time: the Vietnamese forest management style was centralised and private property rights were not recognized. The means of production, including forests and forestland, were under direct management of the state. This management model, in combination with weak law enforcement and limited local participation was one of the reasons leading to net deforestation during the 1945 – early 1990s period.

In 1986, the country switched its politico-economic management model from a centrally-planned economy to a market-oriented economy through the introduction the so-called *"Doi Moi"*[1] renovation policy. The main characteristic of this new economic management model was the recognition of private ownership, which is a prerequisite for the allocation of forests and forestland to individuals. Currently, private forest owners are one of the biggest groups of forest owners in the country. After the transformation of the management model, the forested area of the country increased significantly. However, the contribution of forestland allocation to this net reforestation is still a topic of debate.

---

[1] *Doi Moi* is a Vietnamese phrase that means renovation or reform.

The current Vietnamese forestry policies stress the importance of free-market mechanisms. The central theme of the policies is to balance economic development with environmental protection and conservation. It is believed that allocating forestland to individuals and using free-market mechanisms can create economic incentives to promote the efficient use of privately-owned resources.

However, the effectiveness of policies promoting the efficient use of privately-owned resources depends on the understanding of forest owners' decision behaviour in planting and harvesting trees as well as how non-industrial private forest (NIPF) owners respond to changes in social, economic and institutional conditions (van Putten and Jennings 2010).

The decision behaviours of NIPF owners has been posited to be more complex than their industrial counterparts due to the wide range of objectives of the owners (Gregory et al. 2003). Different forest owners have their own objectives that shape their activities. The owners may also adjust their forest management objectives due to changes in socio-economic development (Hengeveld et al. 2017). Thus, it is important to document their decision-making processes with the purpose of adapting the existing policies (Côté et al. 2017).

The behaviour of NIPF owners can be divided into two important aspects: behaviours related to planting and harvesting trees. These behaviours are important because they will affect to the future of forest cover and the supply of forest products in the country. Thus, understanding and predicting these behaviours is an increasingly important task in forest management and planning.

To date, it is unclear whether or to what extent the NIPF forest owners' decision behaviours in planting and harvesting trees can be predicted by means of normative models (Musshoff and Maart-Noelck 2014). To understand the forest owners' behaviour, one needs to know what they are, how and why they are formed and sustained, and what consequences they generate.

Taking the points presented above into consideration, this research titled "Understanding factors influencing non-industrial private forest owners' decisions in planting and harvesting trees: a case study in Thai Nguyen province, Vietnam" is developed with the following goals:

1. Identifying factors affecting the tree planting and harvesting decisions of NIPF owners.
2. Quantifying the importance of the factors affecting the forest owners' decisions.

3. Testing different regression approaches in modelling and quantifying the importance of factors affecting the decisions of the forest owners.
4. Modelling the afforestation and harvest intensity of NIPF owners.
5. Developing policy recommendations for forest managers in Vietnam.

This thesis is composed of six themed chapters. Chapter two details the evolution of the forestry landscape in Vietnam with the purpose of framing the significance of this research, as well as reviewing similar research on this topic. Chapter three discusses the nature of policy, the conceptual framework of the research and presents a brief introduction to the region of study. Chapter four presents the design and results of the first study. Chapter five reports the design and results of second study. Chapter six provides general conclusions of the research.

CHAPTER 2: LITERATURE REVIEW

2.1. Introduction

This chapter presents a summary of the historical transformation of forestry governance in Vietnam and its associated impacts on the national forestry sector. This chapter also includes a review of the current literature on the tree planting and harvesting decisions of non-industrial private forest (NIPF) owners. In this review, the term forestry landscape is used to refer collectively to forestland, forests, their biological condition and forestry policies.

The chapter comprises four sections. Section 2.2 presents the evolution of the forestry landscape in Vietnam and its consequences with the purpose of stressing the need for this research. Section 2.3 reviews similar research that was carried out in Vietnam and worldwide. Section 2.4 provides a general summary of the chapter.

2.2. Evolution of the Forestry Landscape in Vietnam

2.2.1. The Centralised Management Period

The history of the Vietnamese forestry management system and its policy transformation can be traced back to 1945 when Vietnamese revolutionary leaders declared independence on 2 September 1945 and announced the creation of the Democratic Republic of Vietnam. The country established a socialist form of state forestry management, which went unaltered by the government of the Socialist Republic of Vietnam after the reunification of the North and South of Vietnam in 1975 (Dang et al. 2012).

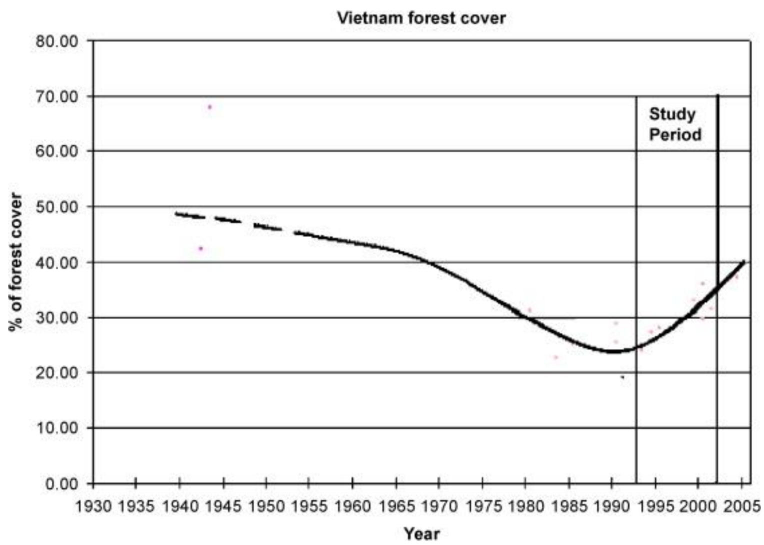


Figure 1: History of Vietnam forest cover (Adapted from Meyfroidt and Lambin (2008))

Figure 1 presents historical information on Vietnam forest cover from 1930 to 2005. As can be seen in Figure 1 from early 1945 to the early 1990s, there was a significant reduction in the national forest cover percentage. It was caused by direct and indirect drivers. The direct drivers were wars in Vietnam, forest fires, legal and illegal logging, and the slash and burn practice (Jakobsen et al. 2007; Khuc et al. 2018; McElwee 2004; Sikor and To 2011; Westing 1971).

The indirect drivers were an increase in population and inefficient agricultural practices resulting in a low rate of production. Consequently, people extended their agricultural fields into forests to boost agricultural production (Dao Minh et al. 2017; Tachibana et al. 2001). The increasing demand for forestry-related products also led to over-exploited forests.

Another notable indirect reason was the political economy of the country. The Vietnamese politico-economic model was centralised planning management. The main characteristic of this management model was the abolition of private ownership. The means of production was collectivised under state management, including the forests and forestland. The forests and forestland were therefore considered to be the property of entire population and have to be managed by the state (Dang et al. 2012; Howard 1998). The state was the single actor that was allowed to decide how the resources were used.

The forests and forestland were assigned to State Forestry Enterprises (SFEs) to be managed on behalf of the government. The forest-management decisions including forest harvesting, processing, planting operations and rehabilitation were heavily dependent on the instructions from the central government. During this period, SFEs mainly focused on forest harvesting rather than forest protection (McElwee 2004). SFEs were assigned large areas of forests and forestland but they did not have the manpower to manage them. In addition, law enforcement and local participation were weak (Meyfroidt and Lambin 2008; Nguyen et al. 2010). Consequently, the forests were de jure state property but de facto open access.

2.2.2. Decentralised Management Period

From the early 1990s to the early 2000s, the forested area recovered significantly along with a remarkable transformation of the country's politico-economic model (Dang et al. 2018). In 1986, the country transformed its economic management model through the introduction of the *"Doi Moi"* policy. The introduction of this policy shifted the country's economy from a command-and-control to a market-oriented economy. The primary component of this policy was to legalise private ownership and redistribute the means of production to individuals.

In 1988, the introduction of the Land Law enabled the allocation of land to households and individuals (Howard 1998). It also served as a backdrop for forestry reform (Dang et al. 2012). Forests and forestland were redistributed to the local populace, which now is one of the largest owner groups in the country (Dinh et al. 2017). This process is often referred to in the literature as Forest Land Allocation (FLA). The primary purpose of FLA was to create economic incentives for the local populace to manage their forests and forestland in the proper manner (Nguyen 2006).

However, the contribution of FLA to the observed increase in forest cover in Vietnam still is a subject for debate. Clement and Amezaga (2009) argued that FLA had little impact on a household’s tree-planting decisions. This argument was supported by Nguyen and Masuda (2018), noting that even though forestland recipients planted trees, the actual planted area was less than the allocated land. It can be said that the forestland usage of FLA recipients is currently not well-understood.

During this period, various governmental tree - planting subsidy programs were implemented such as Program 327 and the 5 million ha program that were carried out in 1993 and 1998 respectively (Howard 1998). The goals of these programs were to increase forest cover, encourage the effective use of forestland, increase the income of the people living in rural areas and in mountainous regions, and protect the environment. Dao Minh et al. (2017) and Howard (1998) claimed that the increase in forested area was due to these governmental tree-planting subsidy programs.

Castella et al. (2006) and Dang et al. (2012) argued that the increase in forest cover was caused by agricultural intensification and the development of the agricultural market in Vietnam, which helped to reduce the pressure on forest owners to use forestland for agricultural purposes.

2.2.3. Current Narratives

The current forestry policies in Vietnam are a legacy of the transformation of the national economic model from a centralised command economy to a market-oriented economy and echo current global trends in forest policy-making with regard to climate change. The current theme of forestry management policies is to emphasise the importance of the market mechanism (Khuc et al. 2018; Trædal et al. 2016). The policies try to balance economic development with environmental protection and conservation.

In 2008 the government experimented with a national programme of Payment for Forest Environmental Services (PFES). The underlying reason for this policy implementation was that the previous forestry policies failed to take account and value the environmental services generated by forests and an economic instrument was required to ensure a more systematic approach (Do et al. 2018). In 2010, the policy officially became national policy by the issue of Decree No. 99 of the government.

Another market-based mechanism being implemented in Vietnam is REDD+. The current activities of REDD+ include capacity building, designing a benefit-distribution system for distributing funds from forest environmental services buyers to providers as well as setting up a national system for Measuring, Reporting and Vertification (MRV) (Trædal et al. 2016)

Current research in forestry is also performed in this framework. The current literature often focuses on aspects of forestry governance:

- Analyzing the weaknesses of institutional arrangement (McElwee 2012; Nguyen et al. 2018; Suhardiman et al. 2013).
- Benefit distribution in PFES schemes (Duong and de Groot 2018; Hoang et al. 2013; Pham et al. 2015).
- Social safeguards (Nielsen et al. 2018).

Admittedly, the institutional arrangement plays an important role in governance, because the existence of state institutions is not only to secure but also improve social welfare. Hence, a well- structured governmental arrangement can help to sustainably manage forestry resources. Research on benefit distributions in a PFES scheme also is critically important. A well-designed distribution system can attract more people to participate in the system.

However, the focus on institutional arrangement and distribution schemes, to some extent, forms an incomplete picture from a policy standpoint. As noted by Irimie and Essmann (2009), this is a two-way relationship because the targeted economic actors can respond to or alter the institutions that no longer serve their interests.

Hence, it can be argued that the more important aspect of forestry management is to understand forest owner behaviour, especially decisions in planting and harvesting trees. They are key players whose decisions will affect the entire forestry industry in both the short and long term. It would therefore be useful for forestry policy makers to be equipped with a management tool that can help them analyse and understand the decision-making process of private forest owners by using their in-house management information. The understanding of the forest owner decision-making process requires a deep understanding of economic incentives and the factors affecting their forests and forestland management decisions.

However, there has been little discussion on this topic in Vietnam. As such, research into non-industrial private forest (NIPF) owners' decisions in planting and harvesting trees in Vietnam would provide useful inputs for forestry policy making and management. The results of the research can be used by Vietnamese forestry managers as a tool for analysing and developing their forestry policies in the future.

2.2.4. Plantation Forestry in Vietnam

This section introduces the current status of plantation forestry in Vietnam with the purpose of providing meaningful background to the research as well as highlighting impacts of Vietnam socio-economic transformation on plantation forestry.

The first impact of the transformation is the increase in plantation area of the country. According to Ministry of Agriculture and Rural Development (MARD) (2017), in 2016 Vietnam has 4.1[2] million ha of formal plantation of which 2.8 million ha is production forests. There is 631 000 ha of informal plantation that were privately planted on agricultural and other non-forest lands, around houses, roads and streams. The majority of the plantation area is planted in acacia. Midgley et al. (2017b) mentioned that the informal plantation makes considerable contribution to Vietnam's supplies of commercial wood. According to Harwood et al. (2017) Vietnam has over 1.2 million ha of acacia plantation of which clonal acacia hybrid (*Acacia mangium x A.auriculiformis*) accounts for over 500 000 ha.

Secondly, due to recognition of private ownership, private forest owners have become the largest groups of owners in the country, referred here as NIPF owners or smallholders. According to MARD (2017) 1.7 million ha of plantation is managed by individuals and households. Meanwhile, state-owned forestry enterprises currently own about 10% of plantation area. Similar to other South-East Asian countries, this group of forest owners generally own 2 ha and often less than 0.5 ha (Flanagan et al. 2019). They live in crowded rural area and are poor according to Western standards. Therefore, it limits the potential for developing large contiguous commercial plantation. However, research conducted by Frey et

[2] Number was rounded by researcher.

al. (2018) pointed out that the investment return of smallholders is profitable and has a higher rate of return in comparison with state-owned forestry enterprises.

The transformation of the national economic model leads to the flourishing development of a market for trading forest products and made Vietnam become a world-class wood exporter. Vietnam's wood was exported to 120 countries and has risen to more than US$7 billion in 2015 Hoai (2015, as cited in Midgley et al. 2017a). This has made the country the world's largest exporter of hardwood chips and the world's fourth largest exporter of wooden furniture.

Another reason leading to the thriving development of the forest product market is that the government has eliminated unnecessary regulations that may prevent farmers participating in the market and has also helped to reduce transaction cost. Vietnam is an exceptional case in Asia in that smallholders are largely free from constraints with respect to harvesting and transporting forest products to and beyond the farm gate Byron (2017, as cited in Midgley et al. 2017a).

In many Asian countries, most smallholder tree-farmers are severely disadvantaged by complicated sets of government laws and regulations which are implemented at the local level by forest officers (Midgley et al. 2017b). A large proportion of smallholders in Southeast Asia are poorly educated and unfamiliar with the legal demand for wood products (Flanagan et al. 2019). Therefore, compliance with regulations normally involves a considerable transaction cost at many stages of harvesting and transporting forest products to and beyond the farm gate.

Additionally, the presence of traders or forest harvest contractors during the forest harvesting stage plays an important role in the development of the supply chain. The traders are not just timber brokers but also facilitators who efficiently link growers with processors (Flanagan et al. 2019). The tree farmers normally do not know the actual volume of their standing forests. The traders simply offer a price for the whole standing forest. In negotiation with the traders, the farmers may consider the price paid to surrounding smallholders as well as their forest and tree size in comparison with their surrounding forests. The total financial return of forests may heavily depend on this negotiation process.

Along with the development of a market for trading tangible forest products, the Vietnamese government recently focused on developing a market for trading intangible forest products such as environmental services. In 2017, the Vietnam National Assembly issued a Forestry

Law replacing Law on Forest Protection and Development issued in 2004. Two new terms are introduced into this law. These are (i) leasing forest environment and (ii) payment for forest environmental services. These points are the fundamental foundation for developing a market for trading intangible forest products in the future. However, the success and effect of this policy implementation on the rate of reforestation and total wood supply of the country are still topics of debate as mentioned in the previous section.

Last but not least, the success of plantation forestry in Vietnam in recent years has been due to the contribution of advanced knowledge and technologies provided by experts from Vietnam and worldwide. In the 1980s and 1990s, programs led to genetic improvement of plantation trees including seed source trials, traditional and hybrid breeding and clonal propagation (Bartlett 2016; Harwood et al. 2015)

2.3. Literature Review

Decisions are made when a person is faced with a choice between more than one course of action (Lindley 1971; Mankiw 2018; Ostrom 2005). Therefore, it can be understood that any decision made by NIPF owners represents a course of action taken in a certain situation and that the final decision is typically assumed to be a result of considering all relevant information in the decision-making environment (Ficko and Boncina 2013).

In this review, each decision is mathematically considered to be a function of several relevant factors. The output of the function is the decision in tree planting (TPLNT) and tree harvesting (HVRT). The inputs of the function are factors affecting their decisions in planting and harvesting trees. For convenience these factors are grouped into three primary categories: Owner Characteristics (OC), Institutional Factors (IF) and Market Conditions (MC).

The decisions in tree planting (TPLNT) and tree harvesting (HVRT) is presented as a function of the aforementioned groups of variables in the general equation below:

Equation 1: General Form of Tree Planting and Harvesting Model

$$TPLNT/HVRT = f\,(OC, IF, MC)$$

Where:

- The Owner Characteristics (OC) category includes demographic information, resources and preferences of the owners. Demographic information consists of variables such as age, gender, level of education and family structure. The resources consist of variables such as number of forestland plots, plot size and plot soil conditions. The preferences include the forest and forestland management objectives of the owners.

- The Institutional Factors (IF) category is used to indicate governmental factors affecting the owner’s decisions. These include tree planting subsidy programs and technical support from extension workers.
- The Market Conditions (MC) category includes variables such as timber price, cost of buying seedlings and the cost of harvesting.

2.3.1. Tree Planting

2.3.1.1. Owner Characteristics

The owner’s characteristics are the most common variables included in tree planting models. One common choice is to use the age of the forest owners as a predictor. Beach et al. (2005) in his review of twelve studies in USA found that age was not significant in the models. The research conducted by Kulindwa (2016) in Tanzania and Sikor and Baggio (2014) and Dinh et al. (2017) in Vietnam also found similar results.

Research conducted by Ruseva et al. (2015) in USA and Meijer et al. (2015) in Malawi also attempted to use gender, level of education and the family size of the forest owners as predictors. However, only Ruseva et al. (2015) found that the level of education positively correlated with tree planting decisions in the USA.

Level of income is another common choice of predictor for tree planting. It is considered as a measure of the forest owner’s available resources for planting trees, given that tree planting requires upfront investment such as buying seedlings and site preparation. In the review by Beach et al. (2005), it was shown that the level of income was positively correlated with tree planting because it implies better access to capital for tree planting.

Research conducted by Sikor and Baggio (2014) in a northern and southern province in Vietnam also demonstrated that more affluent households tended to have larger forestland plots and invested bigger sums than poorer ones. However, the authors also admitted that the poor households still can participate in tree planting by using cheap seedlings, little use of fertilizers and using household’s labour with the purpose of reducing investment cost.

Dinh et al. (2017) added that the contribution of off-farm and agricultural income to annual income was negatively correlated with tree-planting decisions. It is logical because the greater contribution of these sources of income to annual income implies that the forest owners spend less time in tree planting.

In theory, forestry production processes might depend heavily on plot conditions where the forestry production process occurs. These conditions include land quality and plot size. Soil quality has often been considered as being important factors affecting afforestation decisions. Research conducted by Boulay et al. (2012) in Thailand indicates that forest owners who has sandy soil were more likely to adopt eucalyptus because the forest owners considered eucalyptus the only viable crop for these plots.

Research conducted by Ruseva et al. (2015) in USA shows that small plot landowners tend to manage forests for non-timber products. Additionally, research conducted by Frayer et al. (2014) in China shows that forest owners who own larger areas of crop land tend to plant trees. Boulay et al. (2012) pointed out that the greater the area of forestland owned by forest owners, the greater the adoption of eucalyptus in Thailand.

However, plot size has a mixed effect in the Vietnamese context. Sikor and Baggio (2014) found that this factor does not significantly affect a forest owner’s decision in planting trees in Vietnam. Meanwhile, Tran et al. (2019) found that total forest land has a mixed effect on the tree planting decision. The total planted area increased for owners who own from 0 to 0.5 ha of forestland and owners who own above 2 ha of forestland. Meanwhile, less area was planted for forest owners who own from 0.5 to 2 ha.

Research conducted by Meijer et al. (2015) in Malawi and Duesberg et al. (2014) in Ireland show that owners who had planted trees in the past had more positive attitudes toward tree planting compared to respondents who had not planted trees on their land. Ruseva et al. (2015) showed that forest owners who managed their forestland for recreational purposes tended to plant trees.

2.3.1.2. Institutional Factors

Typical policy instruments are cost and tax subsidies as well as technical assistance. The use of these policy instruments assumes that tree planting requires upfront capital for site preparation and to cover planting costs. Therefore, government support is required, and often expected, to reduce the investment costs of forest owners. It will create economic incentives for farmers participating in plantation programs. For example, Vietnam and Indonesia have implemented policies promoting plantation by providing farmers with discounted credit, technical support and subsidised inputs (Dermawan et al. 2013). Research conducted by

Frayer et al. (2014); Kulindwa (2016); Sikor and Baggio (2014) and Dinh et al. (2017) found that institutional support is positively correlated with tree planting.

However, the influence level of these support systems is different among different types of farmers (Ruseva et al. 2015). The reasons for this difference are diverse. It may be caused by the institutional system of the country. Research conducted in Vietnam has pointed out that the institutional system creates opportunities for local authorities to interpret central government policies in a way that tends to bring benefits for the local elite classes (Clement and Amezaga 2009; Clement et al. 2009).

Another reason leading to the difference between policy intentions and the outcomes is spatial discourse. Dermawan et al. (2013) have pointed out that at the national level, governments introduce policies that must remain relevant across regions. At the regional and local governmental level, decision-making bodies are interested in how the policies fit regional and local contexts. At the household level, farmers make their decisions based on the information that they possess and take courses of action within a range allowed by the local governments.

Colebatch (2002) has come up with an explanation as to why policy failure exists. He sees policy as a two-dimensional model guiding public behaviour. One of the two dimensions of the model, a vertical dimension, interprets policies as a set of rules. The government is considered to be a single entity with absolute power and it chooses courses of action based on assumptions about public behaviour. These courses of action are then dispatched to the subordinate institutions for implementation with the purpose of maximizing social welfare. In this dimension, the success of the policies heavily depends on the ability and capacity of subordinate institutions instructed to carry out the decision.

The second dimension of the model, a horizontal dimension, interprets policies as the structuring of action. Policy implementation in this dimension is an exercise in collective negotiation. It shifts the nature of the policies from pursuing clear objectives to achieving an outcome which is agreeable with all of the relevant participants (Colebatch 2002). The negotiation process in this context can be understood as a process of coming to a common understanding between the governments and the public. If they do not share common points of view, the policies are less likely to be successful. Admittedly, the policies are unique and powerful tools to guide public behaviour. However, they may be biased if the policies and its intentions are regarded as equivalent to actual public behaviours.

2.3.1.3. Market Conditions

The market category is the final group of variables that is often mentioned in the literature. According to Beach et al. (2005), timber prices, planting costs and interest costs are predictors often included in the research in USA. The authors point out that timber and pulpwood prices would be logically expected to be positively correlated with tree-planting decisions. However, 8 of the 12 studies in USA that were reviewed showed that timber and pulpwood prices were negatively correlated with tree planting.

Regarding the short-term interest rate, only five out of the twelve reviewed studies include it in their research. Four out of these five studies found negative correlations with tree-planting decisions. It is caused by the fact that forest investment is a long-term plan. Thus, the increase in the short-term interest rate leads to an increased total cost of planting trees. Therefore, forest owners have fewer incentives to plant trees.

Consequently, forest owners may not be as interested in planting trees and would rather leave the forestland for natural reforestation or convert the forestland for other purposes. However, the use of interest rate as a predictor may not be accurate in every case. As mentioned in section 2.3.1.1, the poor still can participate in planting trees by using cheap seedlings or family labour forces in order to reduce the cost of investment.

2.3.2. Tree Harvesting

2.3.2.1. Owner Characteristics

The forest owners' characteristics are the most frequently-used predictors in forest harvest related studies. For example, in the US and Scandinavia, the harvesting decision depends on the characteristics of forest owners (Musshoff and Maart-Noelck 2014). The predictors in this category that are often mentioned in the research are age, gender, level of education, owners' economic conditions and landowner presence.

Age and gender may have impacts on the forest owners' ability to harvest trees and their harvesting decisions (Novais and Canadas 2010). This is because the harvesting activities of the forest owners require heavy physical workload. Thus, older forest owners are more likely to use outsourced labour to harvest the trees. This corresponds to an increase in the costs of harvesting trees. The aging owners may also decide not to harvest and instead manage their land for wildlife habitat maintenance purposes or recreational activities. Therefore, the landowner's age plays a significant role on whether they will engage in harvesting trees or leave the land for environmental purposes (Joshi and Arano 2009).

This finding might not be well-suited to the Vietnamese context. As mentioned, Vietnamese forest owners could sell their forests to traders. However, age is a criterion worth investigating. As mentioned by Midgley et al. (2017b) many smallholders consider their forest as a "bank account". Trees often are harvested when the family need cash for a wedding, building a new house, medical treatment or education expense. The older forest owners may not want to harvest their forests and may want to keep their forests as a "bank account" and use it for their retirement.

The level of education is claimed to have an impact on harvesting choices but not on the level of harvesting (Størdal et al. 2008). A negative correlation had been found between the number of years of formal education and harvesting behaviour, according to the work of Dennis (1989). He had assumed that the more highly-educated forest owners value forest amenities higher or they have larger incomes than less highly-educated landowners.

Level of income is also a significant predictor of harvesting behaviour. The total income of a forest owner is the sum of forest income and exogenous income. Dennis (1989) showed that harvesting behaviour is influenced by the level of exogenous income. An increase in the level of exogenous income leads to a reduction in the marginal utility of income derived from timber harvesting. However, a study in Tasmania, Australia, conducted by van Putten and Jennings (2010) showed that total income is the only independent variable that can explain past harvesting decisions and future harvesting intentions.

Regarding financial security, debt condition is a strong motivator for harvesting decisions. Conway et al. (2003) have pointed out that debt conditions limit options for landowners. The landowners with higher debt ratios may accept lower prices for their harvest in order to meet financial obligations and have less incentive to bequeath properties to their heirs. They may also have fewer incentives to be involved in non-timber production activities.

Meanwhile, the residency of landowners is a factor suggested by Conway et al. (2003). They found that non-resident forest owners are less likely to harvest than resident landowners. This is because they consider the property as a place for personal enjoyment rather than a serious timber investment. In addition, it can be assumed that they have a relatively high level of financial security due to their non-residency, which indicates that they are able to rely on other sources of income. Therefore, the income from timber harvesting does not have a significant impact on their decision behaviour.

The number of non-resident landowners can be a good indicator that enables the prediction of the timber supply of the NIPF owner sector. This is because an increase in the number of non-resident landowners owning forestland would result in decreased timber supply from the NIPF owner sector in the future. Furthermore, more non-resident NIPF owners may lead to changes in the plot/resource condition of the forestland.

Beach et al. (2005) had shown that plot size is the most common indicator included in harvesting models and is positively correlated with harvesting in USA. The study conducted by Conway et al. (2003) had suggested that landowners with large forest tract sizes are likely to harvest their forests. Meanwhile, small parcel owners are not likely to harvest trees due to difficulties in finding bidders for their forests.

This research lessons might not fit in the Vietnamese plantation forestry context, because Vietnamese forest owners live in very different socio-economic condition context. As mentioned in section 2.2.4, most smallholder tree-farmers in Vietnam generally own 2 ha and often less than 0.5 ha. They live in crowded rural area and are poor according to developed countries' standards. Therefore, assuming that Vietnamese forest owners' behaviour is similar to that in a developed country with respect to plot size may be impractical.

Furthermore, the small sized forestland plots in combination with the government grant scheme means individuals tend to participate in forest related activities because they can earn extra income if they plant small diameter forestry trees on a short rotation (Tran et al. 2019). In addition, the presence of commercially valuable species on forest land has been proven to have a strong positive correlations with harvesting decisions (Dennis 1989).

However, if the plot size is too small, the forest owners may not plant trees because the small forest owners seek to maximise their cash return per unit of labour input (Byron 2001b). Therefore, the return from selling a small forest plots may not attractive enough for them. Because of this, it is useful, to know how plot size affect the decisions of forest owners with respect to harvesting and planting trees.

2.3.2.2. Institutional Factors

Unlike planting studies, few harvesting studies had attempted to include policy as a variable in their harvesting models. The review of Beach et al. (2005) showed that only three out of twelve studies in the US included this variable in their models. However, the results of including policy as a variable were largely inconsistent, with one study finding a positive correlation, the second finding a negative correlation and the third finding no correlation between policy and harvesting decisions. A study conducted by Ní Dhubháin et al. (2010) in Ireland also indicates that the use of educational instruments as policy intervention has a significant positive impact on forest owners' harvesting decisions. The authors additionally claimed that a low level of forestry knowledge might result in incorrect decision-making regarding thinning. In summary, the impacts of policy variables on NIPF owners' decision behaviour in harvesting trees is currently not well-understood.

2.3.2.3. Market Conditions

Regarding market drivers, timber price is the most common variable included in the NIPF harvesting models. The key assumption made is that an increase in timber prices corresponds to an increase in harvesting intentions (Beach et al. 2005; van Putten and Jennings 2010). In the theory, the actual decision-making behaviour regarding harvesting is analysed on the basis of aggregated data variables such as timber price, inflation adjusted cost of harvesting rate, as well as the amount of timber available to harvest.

However, the empirical results are different from the theory. A study conducted by van Putten and Jennings (2010) in Tasmania, Australia showed that increases in pulp prices are unlikely to affect the harvesting intentions of NIPF owners. However, it is likely to increase the harvesting intensity of those that perform harvesting. Beach et al. (2005) did reviews of the behaviour of twelve studies in the USA and found that the influence of prices is significant and positive only in seven of the reviewed studies. One of the studies found a negative impact and the others found that prices have no impact.

The reasons for the differences may be due to the assumptions that the authors made during their research design. Most of the authors use economic theory as a guideline for their research to explain market drivers. One of the key assumptions in the economic model is that decisions are made in a market where full information on incomes, prices and qualities is available and there is no uncertainty (Kooreman and Wunderink 1997). This method has a strong advantage due to the availability of tools to model and analyse the data (Ficko and Boncina 2013).

However, NIPF owners' decision behaviours operate in incomplete markets in which their preferences and characteristics will alter their harvesting decisions (Midgley et al. 2017a; Størdal et al. 2008). In addition, the NIPF owners' decisions may be shaped by informal information channels of social influence (Ruseva et al. 2015). Therefore, if the economic models are employed, the availability of market-related information must first be ensured.

2.4. Summary

This chapter of the thesis has focused on a summary of the historical transformation of forestry governance in Vietnam and its consequences on forestry plantation. It can be said that Vietnam is an exceptional case that has successfully moved from net deforestation to net reforestation and currently is a world-class wood exporter. The governance reform and availability of technologies are key forces accelerating this positive movement.

It can be said that Vietnam currently has four key conditions which were mentioned by Byron (2001a) in order for widespread tree plantation. These are:

1. Land use is secured
2. Technologies are available
3. Farmers have reasonable confidence that they can protect their trees until merchantable age
4. There is an available market for trading forest products.

There remains a need for research and support to secure sustainable wood production from plantations over successive rotations. The key question to both forest owners and managers is how they secure or accelerate wood production in coming years. The easiest answer is "just do the same things that have been done". This answer could be acceptable if the focus is just on the forestry sector itself. Given that land use rights are secured, policy constraints were lifted, technologies and markets are available, this answer is understandable.

However, we can examine a greater picture of the national economy that does not only contain the forestry sector but also industry and service sectors which normally have a greater contribution to national GDP. These sectors often attract a great number of people to work. This may lead to reallocation of labour resources from forest related activities to non-forest related activities.

Additionally, the flourishing development of forestry plantation in Vietnam from the 1990s to early 2000s has happened in a very special context. The forest owners who were the main force contributing to this achievement were born before the 1990s[3]. At that time, there were few industrial zones and the service sector had not been well developed. Furthermore, the transportation system connecting regions was limited. This limited their ability to move from region to region to live and work. Therefore, focusing on the cultivation of their forestland in order to improve their livelihood based on their skills, knowledge, resources and support of government and technical experts was an understandable decision.

The current situation is to some extent different from this previous period of time. More industrial zones were established. Consequently, the demand for the labour resource is increased. The transportation system connecting region with region is improving. Communication and media systems are developed. Because of this, forest owners do not lock themselves within their farm gate or village border anymore. It provides greater opportunities to allocate their labour resources to non-forest related activities. Moreover, the future generation of forest owners who were born since the late 1990s often have a better education level. Theoretically, they have the chance to allocate their labour to non-forest related activities.

This situation leads to a question whether the statement, "just do the same things that we have done", is enough for sustaining wood production from plantations in coming years. Will future forest owners be willing to do similar routines as their parents did given that the socio-economy has changed? Byron (2001b) pointed out that smallholders may seek to maximise the cash return per unit of labour input. Therefore, if there are opportunities providing better cash return per labour unit, for example working in manufacturing factories, taking these options is understandable.

The efforts of government and experts in the field of forestry in Vietnam should be greatly appreciated and recognised. However, managers and researchers should place the forestry sector within a greater picture of national economy in order to understand opportunities and challenges of the forestry sector in the ever-changing socio-economy. To understand these opportunities and challenges, Byron (2001b) suggested that first we should understand forest owners before considering the role of trees that they may or may not have in their livelihoods.

[3] Assuming that people can start planting trees at the age of 15.

However, there has been little discussion on this topic in Vietnam. As such, research into non-industrial private forest (NIPF) owners' decisions in planting and harvesting trees in Vietnam would provide useful inputs for forestry policy making and management. The results of the research can be used by Vietnamese forestry managers as a tool for analysing and developing their forestry policies in the future.

The second part of this chapter has reviewed some of the studies that have been done globally and nationally on this topic. It is well understood that this topic has been receiving lots of attention from scientists worldwide. The three aforementioned categories of factors that affect planting and harvesting behaviours: owner's characteristics, institutional factors and market conditions are different from case to case as well as in different planting and harvesting models. Moreover, most of the studies tend to focus on a single aspect of the topic such as planting or harvesting behaviour.

Therefore, additional research in planting and harvesting decision behaviour in Vietnam would be useful. The research would not only contribute to scientific knowledge but also provide science-based knowledge to aid Vietnamese forest policy development in the future. According to Colebatch (2002), the key idea of the governance process is to maximize social welfare by guiding public behaviour in a way that brings about greater happiness for citizens. However, to design good policies, power-holding authorities need tools to analyse public behaviour as well as understand factors influencing public behaviour.

Admittedly, some of literature mentioned in this part of the thesis are from developed countries where forest owners live in very different biological and socio-economic conditions. Therefore, their decision-making process might be different and little relevant to Vietnam context. However, reviewing this literature to some extent provides useful lessons for developing this study. It assists in understanding methodologies and frameworks that were used in this type of study. These are important lessons that will be mobilised in later stage of this study. Additional lesson is that when comparing these studies, it can be realised that the decision-making process is diverse and contextually based. Therefore, research should be designed based on the local context and what can be implemented from the study.

From this point of view, this research has been developed with the purposes of providing more knowledge on the decision behaviours of NIPF owners by combining planting and harvesting in an intensive study as well as creating tools for forestry policy intervention in the future.

CHAPTER 3: RESEARCH DESIGN

3.1. Introduction

Chapter 2 discussed the transformation of forest management in Vietnam and also indicated the need for developing instruments to help policy makers design their forestry policies. The chapter also reviewed studies related to decision-making in tree planting and harvesting, demonstrating that the factors affecting decision-making behaviour are diverse and represent different aspects such as personal characteristics, management objectives and market conditions. These behaviours are often inextricably interlinked with policy.

As such, it is essential that one should know what policy is, why and how policy can guide public behaviour, how public behaviour is formed, and lastly how to bring all these points into a single coherent system in order to fully understand the complete forest management framework. Such a system also helps to set the focal level as well as a framework for designing and analysing the research.

This chapter discusses the nature of policy, the conceptual framework of the research and presents a brief introduction to the region of study. Section 3.2 provides a brief discussion about the nature of policy. Section 3.3 introduces the conceptual research framework, links the framework to the context of study and describes the structure of the study. Section 3.4 briefly introduces the region of study. Section 3.5 provides a general summary of the chapter.

3.2. What is Policy?

Policy refers to the expression of decisions which are made by decision-making bodies. Governing bodies identify potential problems that require collective action to solve, make decisions to attempt to solve them, and then use their resources to implement these decisions (Knox 2012). Colebatch (2002) described that policy is usually based on three assumptions about social order: *instrumentality, hierarchy and coherence*.

The first assumption, *instrumentality*, refers to the purpose of organisations. Organisations exist to achieve objectives in particular areas. Policy can therefore be seen as the authoritative determination of what should be done in order to achieve these particular objectives. These objectives may be broadly stated or be more specific. Additionally, objectives may not be immutable and can change over time as the environment evolves. When new factors emerge, organisations would have to adjust their objectives in order to adapt to new conditions.

For example, the original REDD (Reducing Emissions from Deforestation and forest Degradation) policy focused solely on reducing emissions from deforestation and forest degradation. In 2010, new objective components were added to the REDD policy, and its name was modified to REDD+. These additional components are: the conservation of forests, the sustainable management of forests and the enhancement of forest carbon stock (theREDDdesk 2016).

The second assumption about policy is *hierarchy,* which implies that each policy's procedures secure the endorsement of a single course of action. The policy needs a hierarchical structure to ensure that it maintains its original goals from decision-makers who wield authority to the affected audiences.

The last assumption is *coherence,* which is the fitting of all actions together and the forming of a single system. It is also about how the system should be steered. At the decision-making level, the decision-makers with power want to achieve their goals that are stated in the policies. Meanwhile, on the ground the participants shape their activities in ways that reflect their own interests or points of view. Because of this, the decision makers should choose actions that are more like patterns of interconnection between different participants such as production, markets and policies.

On the other hand, the authorities should maintain their normative power. All things being considered, the policy activities should be about coordination in an attempt to change what the targeted audiences do. For example, the Ministry for Primary Industries of New Zealand cannot go to the field to plant trees on behalf of the forest owners. Instead, they use policies as instruments to facilitate tree-planting actions.

In summary, policies can be understood as the instruments that organisations use to achieve particular goals by mobilising their resources. The policies should address the patterns of interactions of the participants, transforming into courses of actions that operate through a hierarchical structure. Consequently, the decision-makers should realise the pattern of interactions and decide which courses of action should be implemented. This leads to a critical question: how do we know whether the actions are appropriate? One useful way of determining the appropriateness of actions is through the principle of utility.

A notable British philosopher, Jeremy Bentham, who was founder of modern utilitarianism, defined morality in terms of the principle of utility. He argued that no human's actions are possible without motive. The best moral motive may be an action in which a person pursues the greatest happiness of all those affected even at the expense of his own happiness (Parekh 2011). In terms of public administrative management, the governments are the ones that should bring the greatest happiness to the greatest number of citizens

All things being considered, the existence of the government is not only to keep society stable but also improve the social welfare of its citizens by using societal resources such as the tax contributions of citizens, and the country's natural resources. The policies can be understood as the government's instruments intervening in public behaviour in order to maximise social welfare.

Additionally, the policies can also be understood as an outcome of the problem-solving process in which governments investigate and find the solutions to problems. Governments do not directly solve the problems, but they introduce the policies in order to guide or change the behaviour of the targeted audiences in a manner that is expected to solve the problems. When designing the policies, the most important issue is to be able to understand the behaviours of the targeted citizens.

3.3. Conceptual Framework

The ultimate goal of this research is to understand the behaviour of the NIPF owners with respect to tree planting and harvesting activities. One needs to know factors affecting their decisions, how these factors are formed, and what consequences these factors generate. This is the process of learning what it does, how and why it works, and how to create or modify it. Therefore, it is necessary to have a general framework as a basis for setting the focal level and sampling strategies, analysing data and making judgments, as well as designing policy recommendations based on research findings.

The Institutional Analysis and Development framework (IAD) introduced by Ostrom (2005) is used as the principal framework guiding direction of this research. The framework is represented in Figure 2.

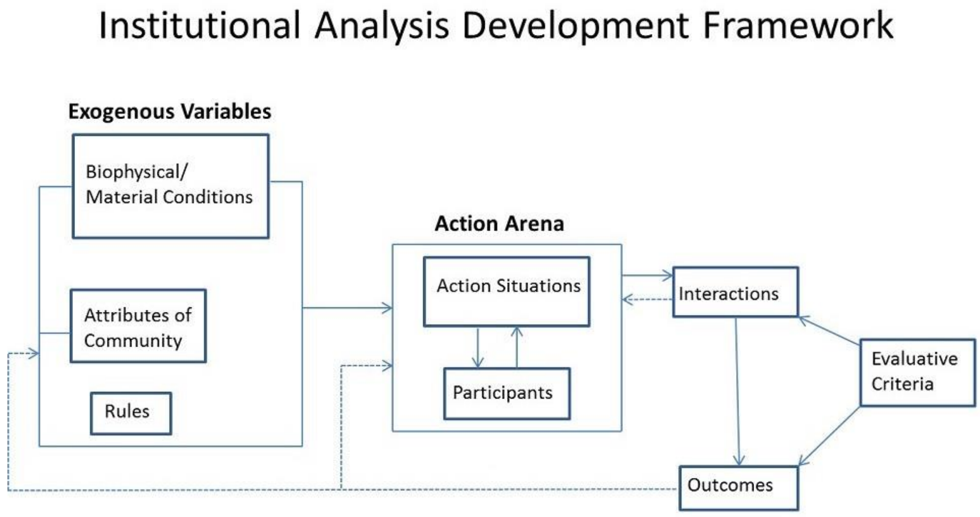


Figure 2: Institutional Analysis Development framework (Adapted from Ostrom (2005), p.15)

In the framework, the exogenous variables form an action arena structure that generates interactions, which then produces outcomes. Evaluative criteria are employed to judge the performance of the system by evaluating the pattern of interactions and outcomes, which feed back into the action arena. This may transform the action arena structure leading to new patterns of interactions and outcomes. Over time, the outcomes start to have an impact on the exogenous variables, further changing the structure of the action arena.

Exogenous Variables

The exogenous variables comprise three main components: biophysical/material conditions, attributes of community, and rules. The biophysical/material conditions are features creating action situations. These features can be soil conditions, and timber and seedling price, etc.

The attributes of community are values that are generally accepted by the community and culture (Ostrom et al. 1994). The attributes of the community may also be factors indicating the perceptions of the participants about the material conditions. For example, it can be how and why the forest owners manage their forests and forestland.

The rules, in this framework, are understood as results of implicit or explicit efforts to achieve order and predictability among citizens by establishing classes of people who are required, permitted or forbidden to take courses of action associated with required, permitted or forbidden outcomes, with the alternative being a high likelihood of being monitored and sanctioned in a predictable fashion (Ostrom 2005). These rules take various forms such as enterprise rules, community regulations, or national laws and policies. Colebatch (2002) also describes policies as tools that are used by a single entity holding absolute power (e.g. governments) to guide and to orient public behaviour. The laws and the policies guide public behaviour by outlining courses of action that are allowed to be carried out and sanctioning inappropriate behaviour. Therefore, laws and the policies are major factors affecting public behaviour.

In conclusion, the exogenous variables are independent predictors forming the structure of the action arena where decisions are made. The rules are established to guide the behaviour of the participants who live in the specific biophysical/material conditions forming communities with their own attributes.

Action Arena

The action arena consists of action situations and participants. Action situations comprise biophysical/material conditions, attributes of the community and the rules. According to Ostrom et al. (1994) the participants in action situations are the decision-making entities assigned to positions and are capable of choosing actions from a set of alternatives that are made available at the time of occurrence of the decision. The actions are taken in the light of information that is possessed by the participants about how the actions link to outcomes in relation to the costs and benefits of the actions and outcomes. The participants in the action arena are people who have their own preferences, information processing ability and selection criteria as well as resources to carry out their actions. In summary, the action arena is a social space in which the participants interact and exchange goods and services to produce the outcomes.

3.3.1. Linking the Framework with the Research Context

This research uses the IAD framework as the instrument to organize the study. The main goal of this research is to determine the pattern of interaction when the participants are in the position of choosing the actions. The actions are decisions regarding planting and harvesting trees on plots of their private land. In other words, the ultimate goal of this research is to model what and how tree planting and harvesting decisions are made in the action arena or to model the interaction block in the framework. The models will be used as a tool for evaluating the outcome of the policies and the designing of forestry policies in the future by changing the rules in order to achieve the intended goals.

According to the conceptual framework, to model the interaction requires two types of input information: the participants and exogenous variables. In other words, it is necessary to know who the participants or the research subjects are and what the exogenous factors forming the action situations are.

3.3.2. Research Subjects

According to the Law of Forest Protection and Development of Vietnam (2004), forest owners are defined as organizations, households or individuals that are assigned or leased forests or forestland for afforestation and have their forest-use rights as well as their ownership rights over the forests (which can be transferred to other individuals) recognised by the State. Forestry contractors are therefore also considered to be forest owners from a legal standpoint. However, harvest contractors were not included in this study because they are not considered as forest owners.

Additionally, forest plantations and natural forests in Vietnam are classified into three types: Production Forests, Special-Use Forests and Protection Forests. Protection Forests, which are used mainly to protect water sources and topsoil, prevent erosion and desertification, regulate climate and reduce the frequency of natural disasters. Special-Use Forests, which are used mainly for the conservation of nature, contain specimens of the national forest ecosystems and forest biological gene sources that are used for scientific research, the protection of historical and cultural relics as well as landscapes, and serve as spots for recreation and eco-tourism. Production Forests are used primarily for the production and trading of timber and non-timber forest products.

Each type of forest owner has an associated bundle of property-use rights with respect to these three types of forest. These rights include access (to enter the demarcated resource area), exclusion (to determine who can use the resources), withdrawal (to extract the resources), management (to modify the resources) and alienation (to transfer the rights over the resources to others).

Table 1 describes the various bundles of property-use rights of the forest owners in Vietnam. As can be seen from the table, the forest owners of production forests have full rights to exercise their decisions regarding their resources in comparison to others. The owners of production forests can be organisations, households, and individuals.

With respect to households and organisations, their decisions may be the result of a collective decision-making process of others such as family members or company shareholders. Because of this, the research solely focuses on the individuals who own production forests.

and assumes that they are freely capable of choosing their actions without any restrictions from others. In this study, they are referred to as NIPF owners or the forest owners.

Table 1: Structure of ownership rights over the forest (Adapted from Dang et al. (2018), p.765)

Property rights	**Production forests**				**Special-use & protection forests**			
	Forest Owners		Forestry Contractors		Forest Owners		Forestry Contractors	
	FP	NF	FP	NF	FP	NF	FP	NF
Access	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Exclusion	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Withdrawal	Yes	Limited	Agreed with owners	No	Limited	Limited	Limited	No
Management	Yes	Limited	No	No	Limited	Limited	Limited	No
Alienation	Yes	No	No	No	No	No	No	No

FP: Forest Plantations; NF: Natural Forests

3.3.3. Study Structure

Around the time period when this research was designed (2015), there were several important factors that strongly affected the research design process. Firstly, when the research idea was initially established, a complete and detailed dataset containing information about Vietnamese forest owners in Thai Nguyen province was unavailable to access. Therefore, it was regarded as unwise to choose a sample based on the existing information on the forest owners. Using geographical locations as a sample unit was considered as an alternative. The sample was intended to include the locations from different parts of the province.

However, this approach also had downsides and uncertainties with respect to the cost of conducting the study and the quality of collected information. The most important downside was the quality of sample. Critical information for the design of the study was unavailable, such as: (i) how many forest owners were present in the locations; (ii) how many forest owners planted or did not plant trees in the locations, and (iii) how many people would be available and willing to participate in the research. These challenges could lead to an unusable sample. Additionally, if non-ideal locations were selected, the cost of conducting the research would also be increased.

The second factor was the implementation of the Provincial Forest Inventory Project in the province during that time. The purpose of the project was to create a central information dataset about the forest owners in the region. The collected information were identities, ownership types, and forest and forestland descriptions of the land owned by the forest owners. To create this database, technical experts went to villages in the province and organised meetings with the forest owners in each village's common house. As they were tasked with going to the villages in any case, the forest inventory experts were asked to help the researcher to randomly deliver the questionnaire to forest owners during their meetings.

Taking the aforementioned factors into consideration, the research was subdivided into two main studies with different specific goals. The purpose of the first study was to identify exogenous factors affecting decisions in planting and harvesting trees. The second study was an attempt to model the afforestation and tree harvesting decisions of the forest owners. The final products of the second study are models to predict the afforestation and harvesting intensity of forest owners. The detailed design and results of these studies are presented in the following chapters of this thesis.

3.4. Region of Study

3.4.1. Natural Conditions

Thai Nguyen is a mountainous midland province with total natural area of about 353 thousand hectares. The province is located in the north east of Vietnam, and the corresponding geographic coordinates are 21°10' to 22°04' north latitude and 103°58' to 104°45' east longitude. The location and terrain elevation of Thai Nguyen province is presented in Figure 3

Some details on the borders of Thai Nguyen province are as follows:

- The north borders Bac Kan province;
- The south borders Hanoi, the Capital City of Vietnam;
- The west borders Vinh Phuc and Tuyen Quang province;
- The east borders Lang Son and Bac Giang province.

The terrain elevation of Thai Nguyen province is high in the north, north-east and south-west. It lowers to the centre and the South of the province. The north mainly consists of limestone mountains (Vo Nhai district), alternating with sandy terrain in a flat valley suitable for agricultural cultivation. To the southwest is the Tam Dao mountain range with a highest peak of 1,592m with vertical cliffs extending in the direction of northwest – southeast. Vegetation cover plays an important role in regulating water flows and supplying water for agricultural production and the daily subsistence of the local populace.

Thai Nguyen province is in the tropical monsoon climate region. There are two distinct seasons: (i) the hot and humid season, from May to October, and (ii) the dry season from November to April. The average annual rainfall is 1366mm. The average annual air temperature is 23.7^{o}C. The lowest monthly average temperature (in the month of January) is 14.2^{o}C. The highest monthly average temperature is 27.2^{o}C. There are no significant average temperature differences between regions in the province. The number of sunshine hours in the year is about 1,187 hours.

Thai Nguyen has two main rivers: (i) Cong River has a catchment area of about 951km^2. (ii) Cau river has a catchment area of about 3,480km^2. The length of the Cau river flowing through the Thai Nguyen territory is about 110km. Besides these two main rivers, there are numbers of other small rivers flowing through the Thai Nguyen province with total length of about 360km. The hydrological regime in the province depends on two main factors: rainfall and the regulation capacity of the Cong and Cau river catchments. The regime can be divided into two distinct seasons: the flood and dry season. The flood season starts from early May and ends in late October, with the highest floods occurring in June, July, August and September. The water flow in rivers during the season usually accounts for 75% of total annual water volume. The dry season lasts for four months, from December to March. The total water volume in these months is only 1.5 - 2% of the total annual volume of water in all of the rivers in the province.

3.4.2. Socio-economic Conditions

The province is divided into nine administrative units, two cities and seven districts. The province's population is about 1.2 million people. The average population density is 325 people per square kilometre. The highest population density is Thai Nguyen City, with 1,545 people per square kilometre. The lowest population density is in the Vo Nhai District, with 78 people per square kilometre. Figure 4 graphically summarises the key statistics regarding provincial population and labour.

Thai Nguyen province is one of the political and economic centres of the northern mountainous provinces. It is a gateway for socio-economic exchange with the major economic triangle of the capital city of Hanoi – Hai Phong province – Quang Ninh province, and the mountainous midland provinces. Thai Nguyen province is well-known for its transportation infrastructure, possessing large quantities of natural resources for the development of industry, has widespread agroforestry, and boasts high tourist numbers. A large natural area of the province can be used for the forestry sector. Figure 5 graphically presents key statistics of the national account and state budget of the province. It can be seen that agriculture, forestry and fishing contribute a total of 11.6% to the Gross Regional Domestic Product. The forestry sector contributes about 4%.

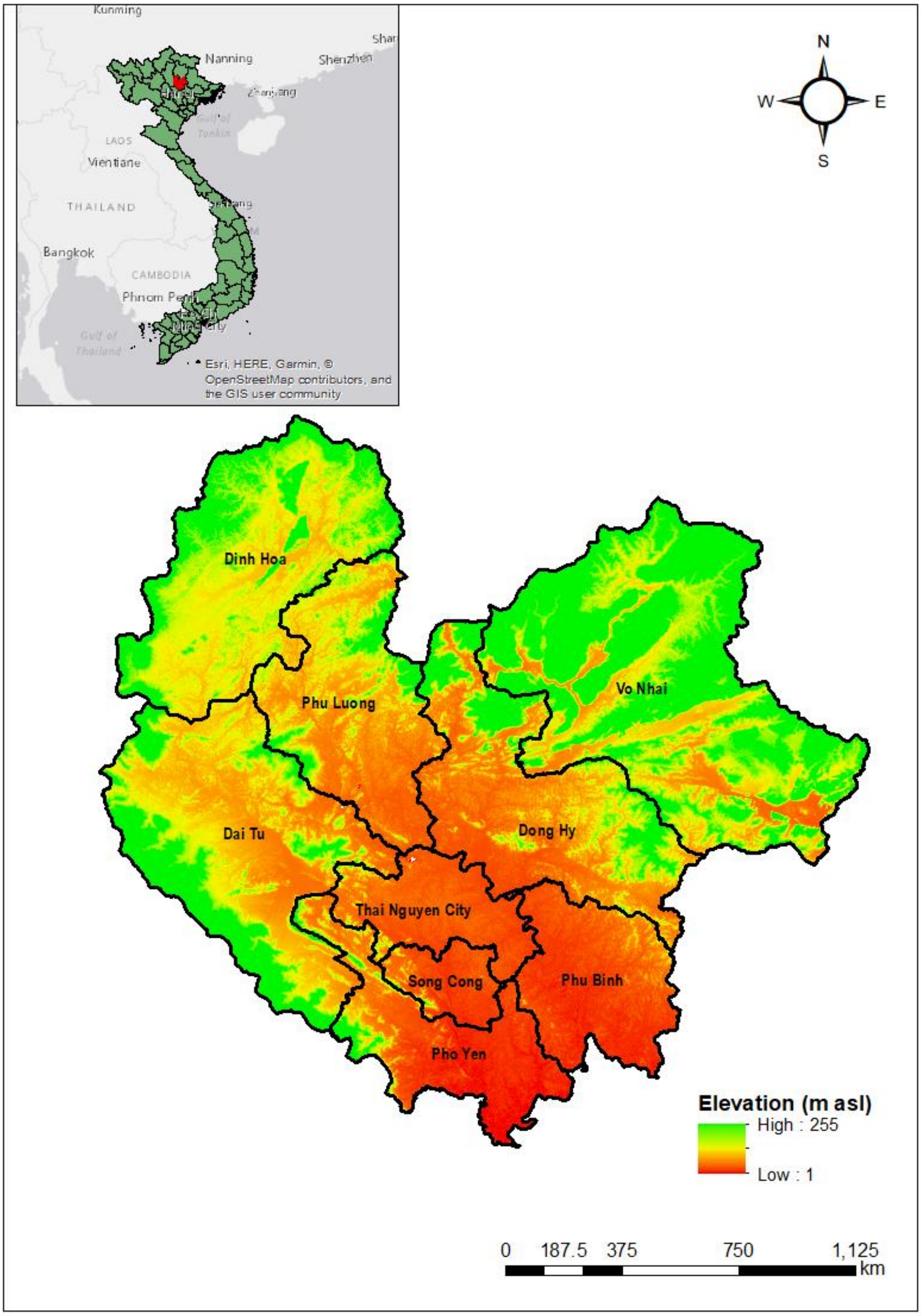


Figure 3: Province Terrain Elevation

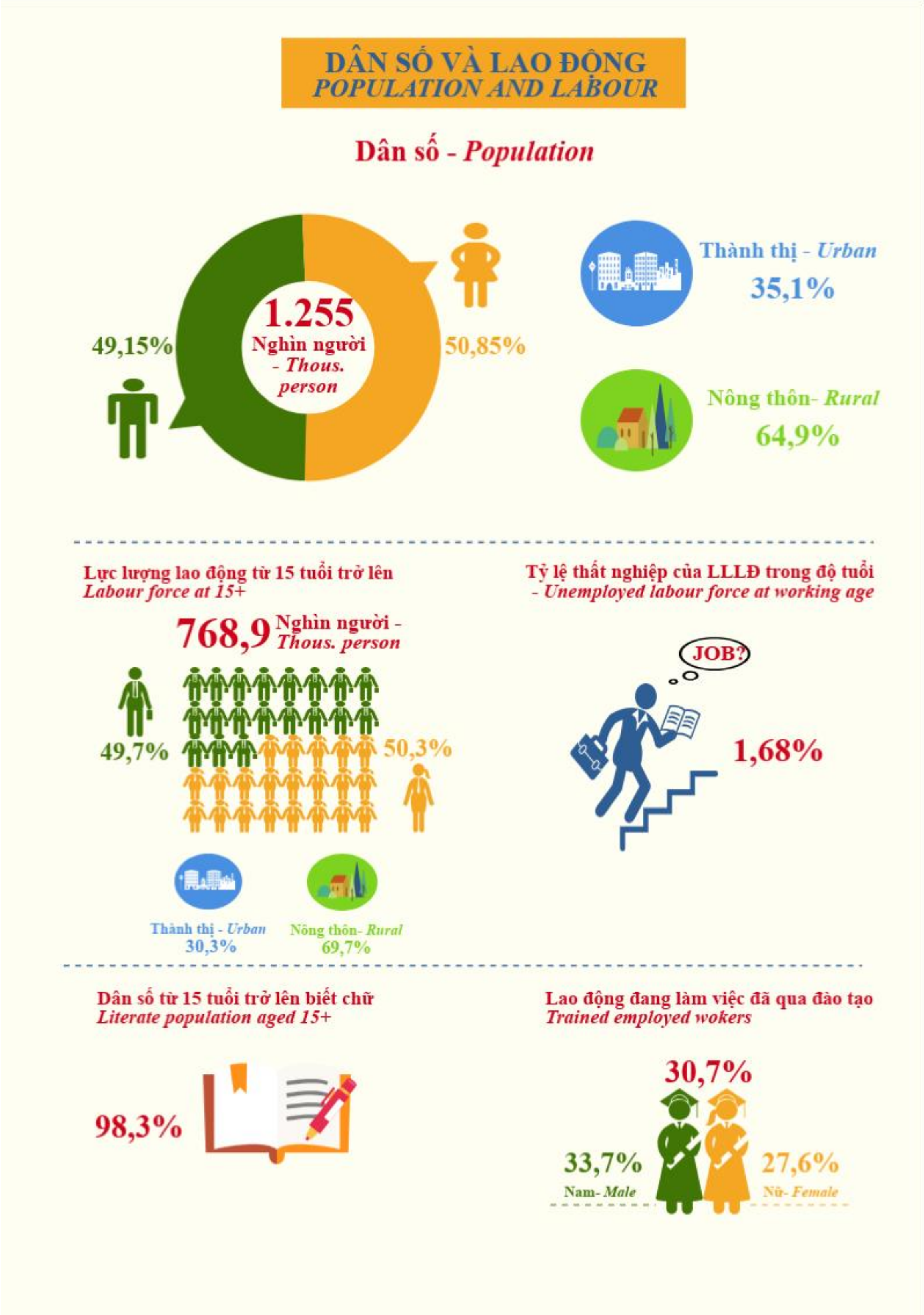


Figure 4: Population and Labour (Adapted from Thai Nguyen Statistic Office (2017), p. 46)

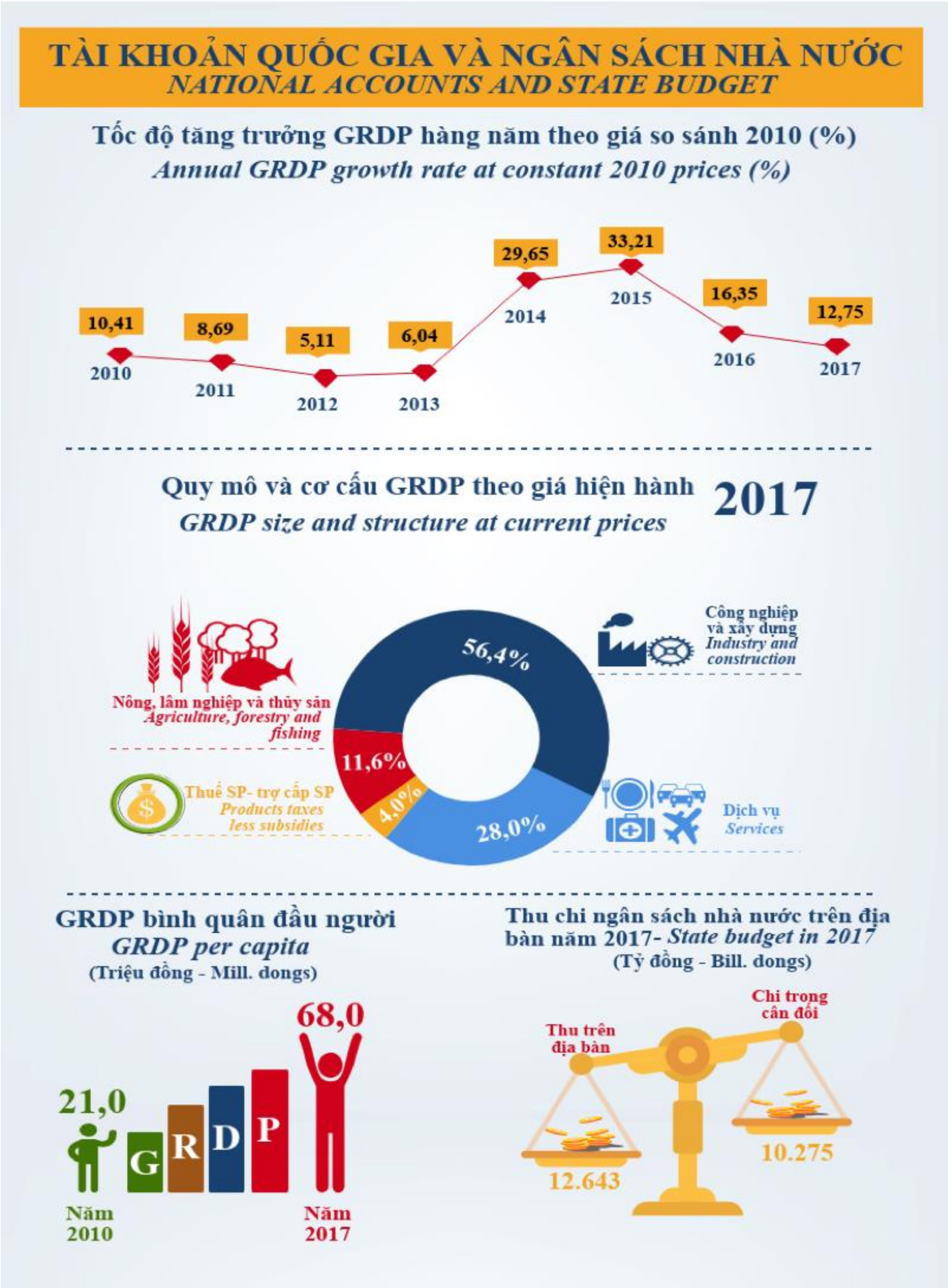


Figure 5: Gross Regional Domestic Product (Adapted from Thai Nguyen Statistic Office (2017), p. 78)

3.4.3. Forest Management

About 180 thousand hectares of the provincial land area, constituting 50% of the total natural land area, is allocated for forestry-use purposes. The allocated areas for special-use forests, protection forests and production forests are 36, 46 and 98 thousand hectares, respectively. Since 1999, about 100 thousand hectares of forestland has been allocated to households and individuals. The remaining areas are being allocated. According to the results of Provincial Forest Inventory Project which was completed in 2016, 28% of allocated forestland area was unplanted. Of these unplanted forestlands, 33% is being using for cultivating agricultural plants but not forestry trees. In addition, the forested areas are typically low-stocked and natural regeneration forests, with few medium- and rich-stocked forests.

The forest management system of the province is hierarchically divided into three levels: the province, district and commune level. The Department of Agriculture and Rural Development is a key governmental institution that advises the Provincial People's Committee in implementing state management with respect to forestry. The management includes establishing forestry plans, planning forest protection and creating development plans, and monitoring forestry activities in the province. At the district level, there are Agricultural and Rural Development Divisions that play a similar role to the Department of Agriculture and Rural Development but advise the District People's Committee instead. At the commune level, the vice-presidents of the Communal People's Committee are in charge for forestry activities in their respective areas. The aforementioned entities are the main decision – makers with respect to their assigned regions.

Along with implementing Law on Forest Protection and Development and other regulations, currently the provincial government has established a well-funded grant scheme for establishing production forests in accordance with the Decision numbered 147/2007/QD-TTg dated in 2007 by the Prime Minister of Vietnam. The grant scheme offers financial and technical support for the forest owners who would like to establish production forests on their own forestland. Financial support can be seedlings and fertilizer or one-time payments to the forest owners if they plant trees using their own seedlings and fertilizer. The provincial government also provides forestry extension workers who would travel personally to the field sites to provide technical guidelines with respect to silviculture techniques.

The grant recipients have absolute rights regarding the established forests including access, withdrawal, exclusion, management, and alienation. When they harvest trees, they must contribute a payment, equivalent to 80 kg of rice, to the Commune and Village Forest Protection and Development Fund. The contribution proportion is 50% to each fund. The funds are not used by the government, but rather for funding forest protection and developmental activities that are carried out by the communities. According to Decision 147/2007/QD-TTg, the financial support of the government can be understood as the payment of the government to the forest owners for environmental services created by their forests during their commercial life cycle.

Currently, *Acacia mangium* is the main species that the provincial government provides to forest owners to establish their production forest. Along with providing seedlings, the provincial government also provides a technical guideline for planting and managing forests. This guideline is similar to what was described by Harwood and Nambiar (2014). There are two main planting seasons. The first season is from March to the middle of June. The second season is from the middle of August to the end of September. Site preparation is manually conducted and the site is burnt before planting. Planting density is 1660 seedlings per ha. Trees are planted in lines with 3m from line to line and 2m from tree to tree.

In the first two years, weed control is by hand and it is conducted twice a year. From year three to harvesting year, little work is required. A typical rotation of acacia plantation in the province is from five to seven years. Normally, when the trees reach the expected harvesting age, the forest owners will look for traders or harvesting contractors or the traders will come to see the forest owners and negotiate with respect to the price of the forest. A better ability to negotiate prices with wood buyers might realise better financial outcomes (Flanagan et al. 2019).

A large proportion of acacia wood will go to local wood chip processors where wood will be chipped then transferred to manufacturing factories or to ports for exporting. The rest of the acacia wood supply goes to local sawmills where it will be used to make products used domestically.

3.5. Summary

This chapter has covered all of the important organisational aspects of the research. The main points of this chapter are:

1. Policy is the authoritative determination of what should be done in order to achieve particular objectives.

2. Policy is based on three main assumptions. These are instrumentality, hierarchy and coherence.
3. The Institutional Analysis Development (IAD) Framework is used as the research conceptual framework. The framework has two main components. These are the exogenous variables and the action arena.
4. The subjects of the research were identified according to their use rights with respect to their forests and forestland.
5. The research is divided into two main studies. The first study focuses on identifying exogenous variables in the IAD framework. The second study focuses on modeling afforestation and harvesting intensity of the forest owners.
6. A large proportion of land area was allocated to households and individuals for forestry-use purposes. However, about 28% of allocated forestland area is unplanted.

CHAPTER 4: IDENTIYFING FACTORS AFFECTING THE FOREST OWNERS' DECISIONS IN PLANTING AND HARVESTING TREES USING REGRESSION APPROACHES

4.1. Introduction

Chapter 3 of this thesis has covered the most important aspects of the research, presenting specifics on how the research is organised. The research is divided into two studies with different objectives and levels of focus. The first study has the following the objectives:

1. Identifying factors affecting the NIPF owners' decisions in planting and harvesting trees.
2. Quantifying the importance of factors affecting decisions in planting and harvesting trees.
3. Testing different regression approaches in modelling and quantifying the importance of factors affecting the decisions of the forest owners.

This chapter of the thesis presents the methodology, results and conclusions of the first study. Section 4.2 introduces the methodology of the study. Section 4.3 describes the process of data gathering. Section 4.4 describes the collected interview forms. Section 4.5 and 4.6 presents data analysis and results of tree planting and harvesting decisions, respectively. Conclusions and limitations of the study are presented in section 4.7 and 4.8, respectively.

4.2. Methodology

Traditionally, a statistical procedure that uses sample data to draw inferences about the population of interest is called hypothesis testing (Gravetter and Wallnau 2009). However, according to Harrell (2015), when developing a statistical predictive model, hypothesis testing and estimation of effects are simply byproducts of the fitted model. Therefore, instead of using a traditional hypothesis testing procedure, statistical predictive models are constructed in this study to identify factors affecting tree planting and harvesting decisions.

The nature of the statistical predictive model is a mathematic function that describes and makes a prediction about a variable of primary interest given a set of explanatory input variables (Harrell 2015; Kruschke 2011). Mathematically, a statistical model can be described as:

Equation 2: A general form of a statical model

$$Y = f(X_j) + e$$

Where:

- Y represents a variable that varies depends on the conditional X.
- X_j j =(1, …, k): k factors affecting the Y.
- e: the residuals that represents factors other than X_j affecting Y.

To develop a predictive model, the variables X and Y as well as a function that describes the relationship between X and Y must be defined.

4.2.1. Variable and Function Nomination

The central research objective of this study is to identify the factors affecting the forest owners' decisions in planting and harvesting trees. Therefore, the response variables are the decisions whether or not to plant/harvest trees. According to Hastie et al. (2009), these types of variables are referred to as discrete variables, which can be numerically represented by codes. A single binary digit, 0 or 1, is often chosen to represent these variables. In the context of this research, these variables are coded as in Table 2:

Table 2: Code rules for response variables

Response Variables	Code Rules
Y_P	0: Never planting trees 1: Planting trees
Y_H	0: Not going to harvest trees 1: Going to harvest trees

The value of the planting variable is based on land description of the forest owners. If the owners planted trees on any plot of their forestland plots, Y_P will be coded as 1. Otherwise, Y_P will be coded as 0. Similarly, with respect to harvesting decisions, if the owners intend to harvest trees, the harvest response Y_H will be coded as 1. Otherwise it is coded as 0.

The nomination of the predictors for the statistical models is based on two main criteria. Firstly, a list of variables is generated according to research lessons that are drawn from literature review in consideration of Vietnam forestry context that is detailed in the literature review chapter.

The second criterion is about the functional use of the final statistical predictive models. The final models are expected to be utilised by Thai Nguyen provincial government in analysing, designing and evaluating their forestry policies' effects at the provincial level by using observational data. Because of that, the predictors should be the ones that can be derived from observational data and are easily interpretable for policy purposes. A list of predictors that aligned with these assumptions were nominated. These predictors are detailed in Appendix 1.

The choice of function depends on the nature of the response as well as the objectives of the model. In this study, the logistic function is chosen to develop the statistical models. The underlying reasons for this choice are presented in Appendix 2.

4.2.2. Identifying Factors

A series of one-predictor logistic models was developed to identify the factors affecting the forest owners' decisions. Other nominated predictors were temporarily considered as unobserved predictors, which are represented by variable e in Equation 2.

Due to the complexity of the data structure and to create reproducible research, it is necessary to have a consistent procedure for analysing the data. The data analysis algorithms were developed with these purposes in mind. Algorithm 1 presents the procedure used for identifying candidate predictors from the set of the nominated predictors. The first specific objective of the first survey was achieved when Algorithm 1 was completed.

Algorithm 1: Candidate predictor identification

The Algorithm 1. Fit the response Y against a single nominated predictor in the set of X_j j = (1, 2, …, j) nominated predictors. 2. Choose k predictors from j nominated predictors that have a p-value of less than 0.1 as candidate predictors. 3. Create a correlation matrix table for these k candidate predictors and calculate the Variance Inflation Factors (VIF).
The Software: 1. R # A statistical environment for running data analysis (R Core Team 2017) 2. Rstudio # An integrated Development Environment for R (RStudio Team 2015) 3. R-package: tidyverse # for data wrangling (Wickham 2017) 4. R-package: usdm # for calculating VIF (Naimi et al. 2014) 5. R-package: SjPlot # For generating correlation matrix (Lüdecke 2018) 6. R-package: stargazer # for generating a model summary (Hlavac 2018)

4.2.3. Quantifying Importance of Factors

The advantage of the aforementioned approach is the reduction of the complexity of the set of nominated variables to a smaller set of the candidate variables. However, this approach presents certain disadvantages. The primary drawback is the difficulty of drawing ceteris paribus conclusions about the impact of individual identified predictors on decisions. This is because the models assume that other investigated factors are unobserved and contained in *e* in Equation 2. This assumption is impractical because the decisions of people are, in fact, more likely to be guided by multiple factors (Ostrom 2005).

Additionally, it is challenging to rank the importance of the predictors if they are only presented in a single-predictor model. Because of that, multiple predictor models are used in order to partially circumvent the aforementioned challenges, as they can be compared and contrasted with each other (James et al. 2013; Mankiw 2018).

This research uses two different regression techniques to develop multiple predictor models: the best subset selection and shrinkage methods. Shrinkage regression employs two different approaches, namely Ridge regression and the Least Absolute Shrinkage and Selection Operator (LASSO).

a) Best subset selection

Best subset selection is an exhaustive screening process that fits and creates a list of models that contain all possible subsets of predictors from the candidate predictors.(Hastie et al. 2009; James et al. 2013). From this list, a set of 100 candidate models are kept for consideration based on a statistical criterion.

An information-theoretical approach is used to make decisions about the choice of models and to rank the importance of the predictors. The nature of the information-theoretical approach can be found in Burnham and Anderson (2002), Burnham et al. (2011) and James et al. (2013).

This research uses the Akaike Information Criterion (AIC) as a mean for choosing the best model and ranking the importance of predictors. The underlying reason for choosing the AIC is presented in Appendix 3. The AIC equations and the method to rank the importance of predictors are presented in Appendix 4. The best subset selection approach follows Algorithm 2.

Algorithm 2: Best subset selection

The Algorithm

1. From k candidate predictors generates all possible combinations of predictors.
2. Fit all those combinations by multiple predictor regression.
3. Keep 100 models with lowest AIC value as the candidate models.
4. Choose the model with lowest AIC value among candidate models as the final model.
5. Quantify the importance of variables using Akaike weights of each model in the set of 100 candidate models.

The Software

1. R # A statistical environment for running data analysis (R Core Team 2017)
2. Rstudio # An Integrated Development Environment for R (RStudio Team 2015)
3. R package for performing best subset selection and multi-model inference (Vincent Calcagno 2013).

b) Shrinkage regression

The best subset selection process as mentioned above can be considered as a discrete process where the predictors are either retained or discarded in the model by an information criterion, for example AIC in this research. Therefore, this method is unable to explicitly illustrate when and how predictors are eliminated from the models. To address the problem, the shrinkage regression, or regularisation, technique is used. A detailed explanation of the technique can be found in Fahrmeir et al. (2013); Hastie et al. (2009); Hastie et al. (2015); James et al. (2013) and Miller (2002).

In brief, the shrinkage method can be understood as a fitting procedure estimating regression coefficients by imposing an additional restriction term (shrinkage penalty) on the values of the coefficients and shrinking them towards zero. The shrinkage regression has different behaviours and names depending on the choice of shrinkage penalty. This study uses two of the most well-known techniques, namely Ridge regression and Least Absolute Shrinkage and Selection Operator (LASSO). Equation 3 and Equation 4 present penalised estimators of Ridge and LASSO regression, respectively.

Equation 3: Ridge regression estimator

$$\text{MLE} + \lambda \textstyle\sum_{i=1}^{p} \beta j^2$$

Equation 4: LASSO estimator

$$\text{MLE} + \lambda \textstyle\sum_{i=1}^{p} |\beta|$$

Where:

- MLE: original Maximum Likelihood Estimator
- *p:* Number of predictors
- $\lambda \sum_{i=1}^{p} \beta j^2$: *L2* shrinkage penalty or Ridge penalty
- $\lambda \sum_{i=1}^{p} |\beta|$: *L1* shrinkage penalty or LASSO
- λ (lambda): A tuning parameter controlling the shrinkage value. If $\lambda = 0$, the shrinkage penalty has no impact and the shrinkage estimator becomes equivalent to the original Maximum Likelihood Estimator. If $\lambda \rightarrow \infty$ all coefficients will be shrunk toward to zero. The model is then determined by the intercepts. The shrinkage method does not apply for the intercept.

A key difference between two techniques is that LASSO creates sparse models as lambda increases but Ridge regression does not. In the LASSO setting, the coefficient estimates of less important predictors are shrunk exactly to zero and eliminated from the model along with the increase in lambda. Meanwhile, Ridge regression does not exclude the predictors. All of the coefficient estimates of the predictors are shrunk by similar factor, λ (lambda). The magnitudes of the coefficient estimates reduce as lambda increases. Consequently, the coefficient estimates of the less important predictors are closer to zero at a given lambda

(James et al. 2013). The shrinkage regression is performed by the glmnet R-package (Friedman et al. 2010)

4.3. Data Gathering Process

4.3.1. Participant Identification

At the time of the designing of this survey, the information about the forest owners in the province was unavailable to use for sampling. Therefore, the study was incorporated into the Thai Nguyen provincial forest inventory project. The main purpose of the inventory project is to create a central database of Thai Nguyen province forest owners. The database includes forest owners' profiles in the province, descriptions of forestland plots, and links the information to digital maps that can be accessed and manipulated through a software interface.

To create this database, technical experts went to villages in the province and organised meetings with the forest owners in each village's common house. The forest inventory information was then collected through fieldwork and notes taken during these meetings. As they were tasked with going to the villages in any case, the experts were asked to help the researcher to randomly deliver the questionnaire to forest owners during their meetings who met certain criteria. These forest owners would have one of the following characteristics:

1) They did not plant trees in their forestland plots
2) They planted trees with assistance from the government tree planting grants.
3) They planted trees without assistance from the government tree planting grant.

4.3.2. Questionnaire Design and Human Ethics Application

This study uses interviews as the main tool for gathering data. These interviews are based on a structured questionnaire, where each question generates a data point for a nominated variable. The questions are divided into four sections in the questionnaire form. The first section collects information about the forest owners' personal characteristics, such as age, family structure, level of education and income. The second section contains questions about the forest management objectives of the forest owners. Information about tree planting and harvesting are collected by the questions in section 3 and 4 respectively.

The order of the sections follows the questionnaire pattern that was suggested by Lavrakas (2008). The first section contains general and neutral questions with purpose of building rapport and obtaining the participant's confidence. Subsequent sections contain questions that require greater effort to address. To maintain confidentiality of the interviewees' data, their names were not recorded.

A separate information sheet was developed with the purpose of introducing the researcher and the nature of the study to the targeted participants. Before starting the interviews, the interviewers read the information sheet to the targeted participants and ensured that they were aware of the interview process and the research. In the end, the targeted participants were offered a chance to refuse the interview, with the implication that if participants participate in the interview, their consent is obtained.

The questionnaire was subsequently sent to the Human Ethics Committee (HEC) of the University of Canterbury for approval. The study was granted approval on 23 September 2015 by the Chair of the HEC.

4.3.3. Translating and Distributing the Interview Forms

The questionnaire and information sheet were translated into Vietnamese by the researcher. The Vietnamese words were carefully chosen to meet the following two criteria:

- Preserving the intent of the sentences as per the English versions.
- Being understandable by people from different educational and cultural backgrounds.

Drafts of the translation were sent to colleagues of the researcher in Vietnam. These people were chosen due to their knowledge and expertise in forestry and experience of working with local people in Vietnam. They were asked to make comments and suggest any required changes to the language style in the drafts. The fundamental goal was to ensure that the final Vietnamese version of the questionnaire was as smooth and natural as if it were initially designed in Vietnamese.

A guideline sheet for distributing interview forms and facilitating the interviews was developed and sent to the forest inventory experts. The guideline sheet contains rules that the respondents and interviewers must follow when asking and answering the questions. A few examples of these guidelines include how to tick an option in the interview form and how to move from one section to another.

4.4. The Collected Interview Forms

The physical interview forms were transformed into electronic ones and subsequently transferred to an SQL database. Before that, a number of collected interview forms were removed from the sample due to the following reasons: (i) incomplete forms due to the participants not completing all of the questions; and (ii) forms collected from incorrect research subjects.

In total, there were 517 usable interview forms, collected from 445 male and 72 female participants. Of these, 289 had completed high school or higher education, while other participants possessed an intermediate school degree or lower level of education. Figure 6 presents the geographical distribution of the participants whose information was recorded in the usable interview forms. It can be seen that the survey covered a great proportion of the province, implying that any conclusions of the subsequent analysis can be considered to be geographically representative.

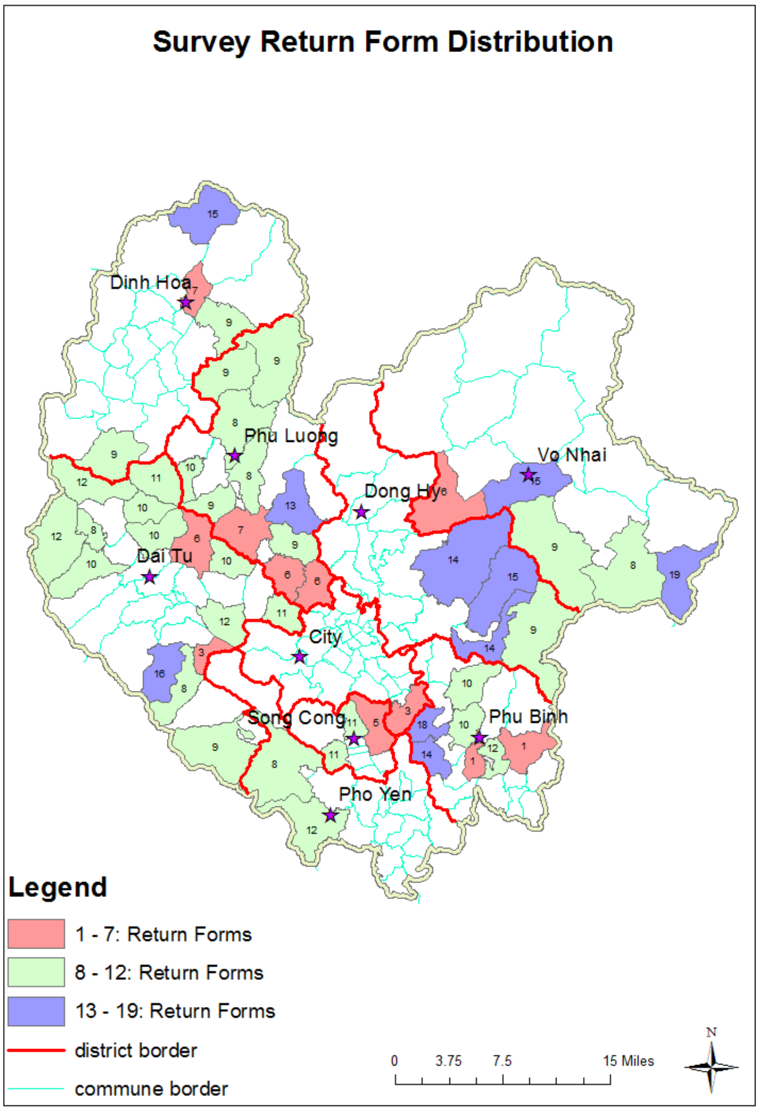


Figure 6: Geographical distribution of usable first study interview forms

Table 3 summarises the key statistical information of the participants in the survey. The average age of the participants is 45. The mean annual income of each participant is approximately 21.1 million VND/year[4]. With respect to family structure, there are an average of five people in each participant's family. Each family has at least one person (over 15 years of age) who is considered to be in the workforce. On average, three people in each participant's family participate in forestry-related jobs. Interestingly, some participants' families do not have any members participating in a forestry-related job.

Each participant owns an average of 1.3 forestland plots with a total forestland area of 0.7 hectares. More than half of the forested plots were planted with *acacia mangium*. The rest of the forested plots were planted with other species including native species. In general, the participants own less cropland than forestland, with each participant owning an average of 0.2 hectares of cropland and some participants not owning any cropland at all. The furthest distance from the participants' houses to their closest forestland plot is 15 kilometres and the mean distance is 1.5 kilometres. Some people have forestland plot next to their house. This is a typical situation in the province where some people live next to their forestland plots and practice agroforestry.

The left and right panel of Figure 7 shows tree planting and harvesting decision-variables, respectively. The left panel shows that 408 participants planted trees on their forestland plots and 109 people never planted trees. There were 242 participants who used the government grant for establishing their forests and 166 people who did not use the grant.

For the harvesting decision-variables, the main factor for classification is the intention to harvest trees in the future. If the participants never planted trees, they do not have trees to harvest. Therefore, 109 participants were excluded from the sample population with regard to harvesting decisions. The sample size for the harvesting model is 408. Of those, 356 people are going to harvest their forests and 52 people had indicated that they are not going to

---

[4] 21.13 million VND is equivalent to 1,352 NZD/year. At the time of writing date the exchange rate is NZD1 = VND 15,622.

Annual income = forestry income + crop income + off-farm income. Forestry income is a one-time payment. Therefore, annual income from forestry income is calculated based on overall payment from harvesting the forests divided by the number of years of the rotation.

harvest their forests. Both groups were asked to explain the reason, age and price of forests that they would like to harvest.

Table 3: Descriptive statistics of the fist survey's participants

No.	***Variables***	***Minimum***	***Mean***	***Standard deviation***	***Maximum***	***Missing values***
1	Age of participants	25	45	11.7	75	0
2	Average annual income of participants (million VND/year)	6	21.1	11.8	46	0
3	Number of people in the participant's house	2	5	1.13	10	0
4	Number of females in the participant's house	1	2	0.9	5	0
5	Number of males in the participant's house	1	3	0.9	6	0
6	Number of people in workforce in the participant's house	1	3	1.2	7	0
7	Number of people participating in forestry-related jobs in the participant's house	0	3	1.2	7	0
8	Distance from the participant's house to closest forestland plot (km)	0	1.5	1.4	15	0
9	Total Forestland Area (ha)	0.01	0.7	1.1	10.2	0
10	Total Number of Forest Parcel	1	1.3	0.6	4	0
11	Crop Land Area (ha)	0	0.2	0.1	1	0

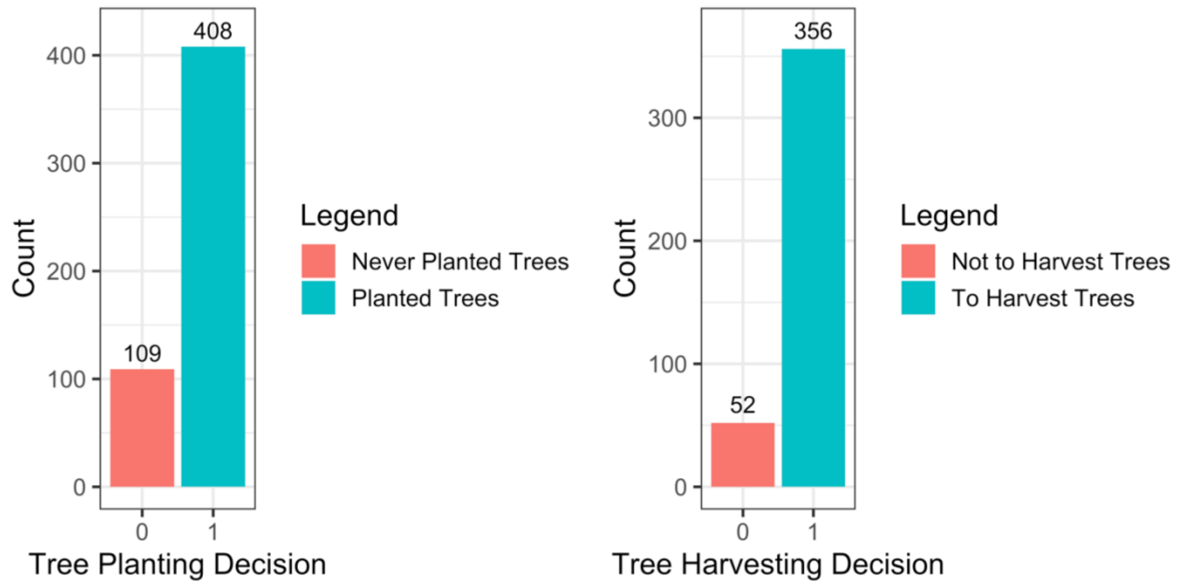


Figure 7: Tree planting and harvesting decisions

4.5. Tree Planting Decisions

4.5.1. Variable Identification

Algorithm 1 is applied to identify candidate predictors affecting decisions in planting trees. The analysis results are presented in Table 4, 5, 6, 7 and 8. Each column in the table is a summary of a model. Each row contains key model information: estimates with standard errors presented in parentheses, the number of observations, the log likelihood and the AIC value of the model.

Table 4 shows a summary of different logistic regressions for tree planting with respect to the forest owners' land asset. The results show that the total cropland area of the forest owner and distance from the forest owners' house to their closest forestland plot were not correlated with the tree-planting decision. The total number of forestland plots and the total forestland area were positively correlated with tree-planting decisions. Therefore, these predictors were chosen as candidate predictors for developing the multiple-predictor model.

Table 4: Summary of tree planting logistic regression with respect to land asset

	CrpA (1)	Parcel (2)	Tfp (3)	Dist (4)
Total cropland area (CrpA)	.73 (1.01)			
Total forestland area (Parcel)		$.64^{***}$ (.23)		
Total forestland plots (Tfp)			$.55^{**}$ (.24)	
Distance from house to closest forestland plot (Dist)				.004 (.08)
Constant	1.20^{***} (.20)	$.99^{***}$ (.15)	$.66^{**}$ (.30)	1.31^{***} (.16)
Observations	517	517	517	517
Log Likelihood	-266.02	-259.94	-263.00	-266.28
Akaike Inf. Crit.	536.04	523.87	530.01	536.57
Note:				$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table 5 presents a summary of logistic models with respect to the forest owners' personal characteristics. The age of the forest owners were negatively correlated with tree-planting decisions. The remaining variables including the gender of the forest owners, the level of education of the forest owners, the number of people living in a house with the forest owners, the number of males and females living in a house with the forest owners, the number of people in workforce living in a house with the forest owners and the number of people who live in a house with the forest owners participating in forestry-related job were not correlated with tree-planting decisions. The age of the forest owners was chosen as candidate predictors for developing the multiple-predictor model.

Table 5: Summary of tree planting logistic regression with respect to personal characteristics

	Age (1)	Gen (2)	edut (3)	Fco (4)	FcoM (5)	FcoF (6)	FcoW (7)	FcoPF (8)
Age	-.02* (.01)							
Gender (Gen)		.02 (.31)						
Education level (edut)			.09 (.22)					
Total number of people in forest owner's family (Fco)				-.06 (.09)				
Total number of males in forest owner's family (FcoM)					-.07 (.12)			
Total number of females in forest owner's family (FcoF)						-.02 (.13)		
Total number of people in the workforce in forest owner's family (FcoW)							-.04 (.09)	
Total number of people particiapting in forestry activities in forest owner's family (FcoPF)								-.04 (.09)
Constant	2.13*** (.44)	1.32*** (.12)	1.27*** (.16)	1.59*** (.47)	1.51*** (.34)	1.37*** (.29)	1.43*** (.31)	1.42*** (.26)
Observations	517	517	517	517	517	517	517	517
Log Likelihood	-264.39	-266.28	-266.20	-266.11	-266.11	-266.27	-266.21	-266.20
Akaike Inf. Crit.	532.78	536.57	536.39	536.22	536.22	536.53	536.41	536.40

Note: *p<0.1; **p<0.05; ***p<0.01

Table 6 presents a summary of logistic models with respect to the forest owners' structure of annual income. The average annual income, the contribution of crop income to annual income and the contribution of off-farm to annual income were negatively correlated with tree planting decision. Only the contribution of forestry income to annual income was positively correlated with tree-planting. These variables were chosen as candidate predictors for developing the final multiple-predictor model.

Table 6: Summary of tree planting logistic regression with respect to structure of annual income

	ConLin (1)	Crpin (2)	Frtin (3)	Othin (4)
Average annual income (ConLin)	-.02* (.01)			
Contribution of crop income to annual income (Crpin)		-.46*** (.17)		
Contribution of foresry income to annual income (Frtin)			.37** (.15)	
Contribution of off-farm income to annual income (Othin)				-.60*** (.18)
Constant	1.66*** (.23)	2.79*** (.55)	.27 (.43)	3.36*** (.63)
Observations	517	517	517	517
Log Likelihood	-264.83	-262.09	-263.22	-260.46
Akaike Inf. Crit.	533.66	528.19	530.44	524.92
Note:		*p<0.1; **p<0.05; ***p<0.01		

Table 7 presents a summary of logistic models with respect to the forest owners forestland management objectives. The management objectives of the forest owners indicate the perception of the forest owners about the forestland. It represents for "Attributes of Community" block in IAD framework that was described in Figure 2. Four forestland management objectives were nominated for investigation. These objectives reflect different strategies of the forest owners for owning/buying forestland. During the interviews, these management objectives were explained as below:

- Reason owning forestland as an investment (own/buy and sell) is to describe a situation that the forest owners own/buy forestland and hope that they can find a buyer who is willing to pay at or more than market value. Cultivating forestland is an optional choice.
- Reason owning forestland for future generation (own/buy and hold) is to describe a situation that the forest owners own/buy forestland and hold it for a period of time. The property will be bequeathed to their heirs in future. Cultivating forestland is an optional choice.
- Reason owning forest land for creating natural landscape (own/buy, cultivate and hold) is to describe a situation that the forest owners are interested in using the actual forestland. The forest owners actively manage and cultivate their forestland and hope to create the aesthetic beauty of landscape.
- Reason owning forestland for generating forestry income (own/buy, cultivate and hold) is to describe a situation that the forest owners are interested in using the actual

forestland. The forest owners actively manage and cultivate their forestland and hope to generate cash income off what the land can produce.

The reasons owning forestland for creating natural landscape, the reason owning forestland for generating forestry income, the frequency of forest-maintenance activities and receiving technical support from forestry extension workers were correlated with tree-planting decisions. These variables were chosen as candidate predictors.

The remaining variables including the reason owning forestland as an investment and reason owning forestland as a way keeping land for future generations were not correlated with tree-planting decisions.

Table 7: Summary of tree planting logistic regression with respect to forestland management objectives

	Roinv (1)	Rokfg (2)	Ronlc (3)	Rogfi (4)	Fmain (5)	Tsup (6)
Reason owning forestland as an investment (Roinv)	.17 (.18)					
Reason owning forestland for future generation (Rokfg)		.21 (.19)				
Reason owning forestland for creating natural landscape (Ronlc)			$.45^{***}$ (.17)			
Reason owning forestland for generating forestry income (Rogfi)				$.55^{***}$ (.16)		
The frequency of forest maintenance activities (Fmain)					1.23^{***} (.18)	
Receiving technical supports from forestry extension workers (Tsup)						1.58^{***} (.23)
Constant	.78 (.57)	.63 (.62)	$.64^{**}$ (.26)	-.31 (.47)	-.13 (.22)	$.54^{***}$ (.14)
Observations	517	517	517	517	517	517
Log Likelihood	-265.83	-265.65	-262.35	-259.92	-239.72	-241.19
Akaike Inf. Crit.	535.66	535.31	528.71	523.84	483.45	486.37

Note: *p<0.1; **p<0.05; ***p<0.01

Table 8 presents the effects of the eight categorical predictors representing factors that the forest owners claimed to have an impact on their decisions. Of these predictors the availability of family members for planting trees, awareness about the government subsidy grant for establishing forests and timber price were correlated with tree-planting decisions. These variables were chosen as candidate predictors for developing the final multiple-predictor model.

The remaining variables including the cost of buying seedlings, source of seedlings in the region, cost of hiring people planting trees, soil condition of the forestland plot, and the size of forestland plot were not correlated with tree-planting decisions.

Table 8: Summary of tree planting logistic regression with respect to other categorical factors

	Bseed (1)	Sseed (2)	Hppl (3)	Scond (4)	FmavP (5)	Govt (6)	TprP (7)	Psize (8)
Cost of buying seedlings (Bseed)	.25 (.22)							
Source of seedlings in the region (Sseed)		-.11 (.22)						
Cost of hiring people planting trees (Hppl)			.23 (.22)					
Soil conditions of forestland plot (Scond)				.26 (.25)				
The availability of family member for planting trees (FmavP)					.42* (.24)			
Awareness about government subsidy grant for establishing forests (Govt)						.75*** (.22)		
Timber price (TprP)							.84*** (.25)	
The size of forestland plot (Psize)								-.27 (.22)
Constant	1.18*** (.16)	1.38*** (.17)	1.20*** (.16)	1.25*** (.13)	1.19*** (.13)	.95*** (.14)	.66*** (.22)	1.44*** (.15)
Observations	517	517	517	517	517	517	517	517
Log Likelihood	-265.65	-266.17	-265.73	-265.72	-264.70	-260.27	-261.18	-265.52
Akaike Inf. Crit.	535.30	536.33	535.46	535.44	533.39	524.54	526.35	535.04

Note: *p<0.1; **p<0.05; ***p<0.01

4.5.2. Variance Inflation Factor (VIF)

Multicollinearity analysis was carried out to exclude predictors that are considered to have high multicollinearity in the set of candidate predictors. The multicollinearity problem was detected by using Variance Inflation Factor (VIF). According to Gujarati and Porter (2009), if the VIF of a predictor exceeds 10, that predictor is considered to be highly multicollinear. Using this prescription, the predictors that had VIF values greater than 10 should be excluded from the set of identified predictors.

The results of the multicollinear analysis are presented in Table 9 and 10. The results suggest that there is no existence of any highly multicollinear problem among 14 candidate predictors. The minimum correlation found was 0.001, which occurred between contribution of crop income to annual income and total forestland plot. The maximum correlation is 0.64, which occurred between reason owning forestland for generating forestry income and contribution of forestry income to annual income.

Table 9: VIF values of tree planting candidate predictors

No.	Predictors	VIF values
1	Total forestland plot (Tfp)	1.33
2	Total forest land area (Parcel)	1.38
3	Average annual income (ConLin)	1.66
4	Contribution of crop income to annual income (Crpin)	1.65
5	Contribution of forestry income to annual income (Frtin)	2.05
6	Contribution of off-farm income to annual income (Othin)	1.15
7	Age of the forest owners (Age)	1.46
8	Reason owning forestland for creating natural landscape (Ronlc)	1.23
9	Reason owning forestland for generating forestry income (Rogfi)	1.97
10	Frequency of forest maintenance activities (Fmain)	1.58
11	Receiving technical support from forestry extension workers (Tsup)	2.17
12	Family member available for planting trees (FmavP)	1.07
13	Awareness about the government subsidy grant for establishing forests (Govt)	1.42
14	Timber price (TprP)	1.35

Table 10: Correlation matrix for tree planting-decision candidate predictors

	Tfp	*Parcel*	*ConLin*	*Crpin*	*Frtin*	*Othin*	*Age*	*Ronlc*	*Rogfi*	*Fmain*	*Tsup*	*FmavP*	*Govt*	*TprP*
Tfp														
Parcel	0.46***													
ConLin	0.19***	0.22***												
Crpin	0.00	-0.09*	-0.37***											
Frtin	0.06	0.13**	-0.15***	0.47***										
Othin	-0.04	-0.10*	0.12**	-0.02	-0.20***									
Age	0.23***	0.24***	0.47***	-0.06	-0.05	0.03								
Ronlc	-0.17***	-0.08	-0.18***	0.01	0.16***	-0.02	-0.28***							
Rogfi	0.02	0.08	-0.21***	0.41***	0.64***	-0.24***	-0.06	0.19***						
Fmain	0.01	0.05	0.12**	-0.22***	-0.00	-0.11*	-0.07	0.08	0.09					
Tsup	-0.03	0.06	-0.11*	-0.08	0.09*	-0.11*	-0.16***	0.23***	0.19***	0.54***				
FmavP	-0.02	-0.02	0.01	-0.03	-0.07	0.04	0.02	-0.04	0.04	-0.03	0.06			
Govt	0.02	0.03	-0.18***	0.17***	0.25***	-0.10*	0.02	0.08	0.25***	0.05	0.24***	0.14**		
TprP	0.13**	0.09*	-0.06	0.23***	0.36***	-0.24***	0.09*	-0.10*	0.33***	-0.03	0.02	0.04	0.31***	

Computed correlation used pearson-method with listwise-deletion.

4.5.3. Best Subset Method

The set of fourteen candidate predictors was used to develop the multiple-predictor models by applying Algorithm 2. The top part of Table 11 presents general information about the set of 100 candidate models. The best and the worst models' AIC are 443 and 446 respectively. The difference between the best and the worst model is 3, implying that no single model is clearly superior to another in the set of candidate models. Because of that any model in the set might provide a reasonable approximation of the tree planting decisions of forest owners.

Due to this situation, it would be unwise to base explanations and predictions on a single model with the best AIC. Therefore, instead of relying on a single model to draw inferences, a multi-model inference framework was used. The multi-model inference framework produces model parameter estimates that are not conditional on a single model but instead derive from the weighted averages of these estimates across multiple models. The mathematical details of this technique can be found in Appendix 4.

The lower part of Table 11 presents a summary of the model parameter estimates under the multi-model inference framework. The summary contains information about the averaged estimates, the number of models in which the predictor appears in the set of 100 candidate models (Nb models), and average importance of predictors. It can be observed that four predictors display negative coefficients in the models. This averaged model correctly predicts 427 out of the 517 cases of tree-planting decisions, or equivalently a 18% misclassification rate. The model terms can be classified into three main groups according to their average level of importance.

1. The first group contains the six most important predictors: receiving technical supports from forestry extension workers, timber price, average annual income, contribution of crop income to annual income, frequency of forest maintenance activities, total forestland area and total forestland plots. These predictors have level of importance above 0.8.
2. The second group includes total forestland plot, family available for planting trees and reason owning forestland for creating natural landscape. These predictors have levels of importance from 0.5 to 0.8.
3. The last group contains the reason owning forestland for generating forestry income, contribution of forestry income to annual income, awareness about government subsidy grant for establishing forests, age of the forest owners and contribution of off-

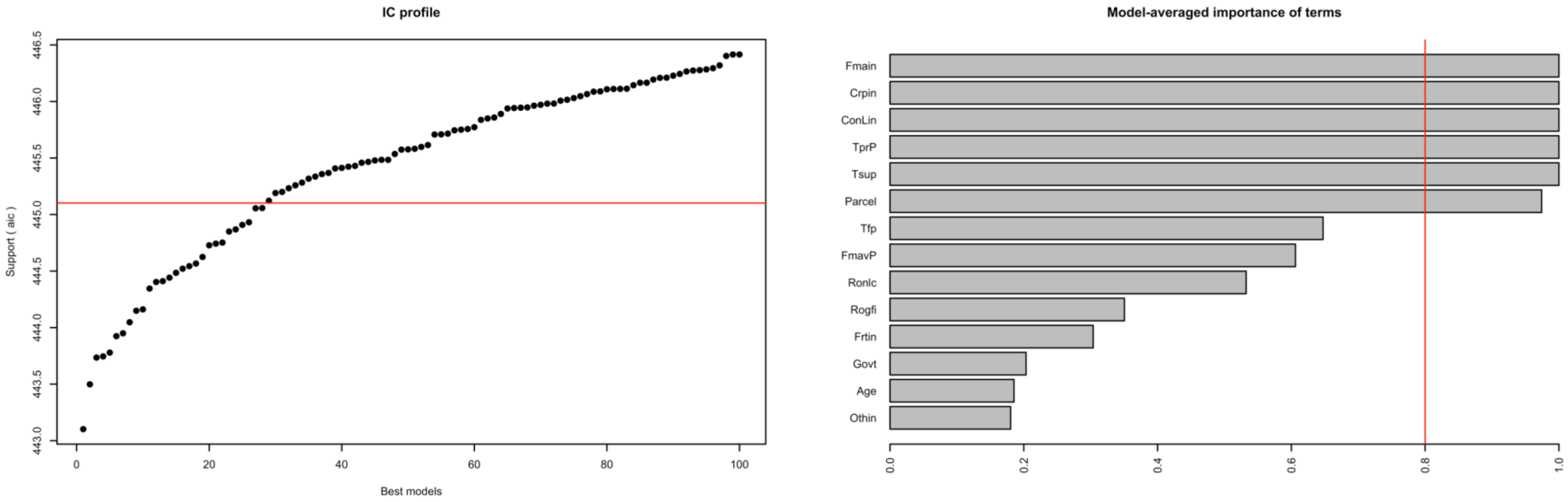
farm income to annual income. These predictors have a level of importance below 0.5.

Figure 8 graphically presents the result of the best subset selection procedure. The left panel presents Information Criterion (IC) profile with the y-axis and x-axis presenting the model's AIC and the model order, respectively. The red horizontal line represents the AIC value that is two units higher than the best model. The right panel in Figure 8 shows the estimated relative importance of the model predictors.

Table 11: Summary of the best subset selection procedure for tree planting decisions

Method: h / Fitting: glm / IC used: AIC
Level: 1 / Marginality: FALSE
From 100 models:
Best IC: 443.1
Worst IC: 446.42

	Averaged Estimates	***Nb models***	***Importance***
(Intercept)	0.893	100	
Receiving technical supports from forestry extension workers (Tsup)	0.804	100	1
Timber price (TprP)	0.924	100	1
Average annual income (ConLin)	-0.036	100	1
Contribution of crop income to annual income (Crpin)	-0.694	100	1
Frequecy of forest maintenance activities (Fmain)	0.782	100	1
Total forestland area (Parcel)	0.555	96	0.974
Total forestland plots (Tfp)	0.308	61	0.647
Family available for planting trees (FmavP)	0.259	59	0.606
Reason owning forestland for creating natural landscape (Ronlc)	0.157	51	0.532
Reason owning forestland for generating forestry income (Rogfi)	0.095	36	0.35
Contribution of forestry income to annual income (Frtin)	0.075	33	0.304
Awareness about government subsidy grant for establishing forests (Govt)	0.04	26	0.203
Age of the forest owners (Age)	-0.001	24	0.185
Contribution of off-farm income to annual income (Othin)	-0.028	22	0.18



Legend:	
Tfp:Total forestland plot	Ronlc: Reason owning forestland for creating natural landscape
Parcel: Total forest land area	Rogfi: Reason owning forestland for generating forestry income
ConLin: Average annual income	Fmain: Frequency of forest maintenance activities
Crpin: Contribution of crop income to annual income	Tsup: Receiving technical support from forestry extension workers
Frtin: Contribution of forestry income to annual income	FmavP: Family member available for planting trees
Othin: Contribution of off-farm income to annual income	Govt: Awareness about the government subsidy grant for establishing forests
Age: Age of the forest owners	TprP: Timber price

Figure 8: Graphical results of the best subset selection for decisions in planting trees.

Left panel: AIC profile. Right panel: Estimated importance of predictors.

4.5.4. Shrinkage Regression

The set of fourteen candidate predictors are used for Ridge and LASSO regression. The left and right panels of Figure 9 graphically present the results of Ridge and LASSO regression, respectively. Each curve corresponds to regression coefficient estimates for one of the predictors, plotted as a function of the natural logarithm of lambda. The lambda values are presented in the top x-axis. The dotted line in the centre of the plot indicates zero-coefficient level. The variables above the dotted line will have positive sign in the model, and vice versa.

At the extreme right-hand side of the plot, lambda is essentially zero, and therefore the corresponding shrinkage coefficient estimates are effectively equal with the usual multiple predictor logistic regression. As lambda increases, the shrinkage coefficient estimates decrease towards zero.

The left panel shows that the Ridge coefficient estimates tend to decrease in aggregate as the natural logarithm of lambda increases. All coefficient estimates are basically zero when the natural logarithm of lambda is greater than 2. The coefficient of the awareness about government subsidy grant for establishing forest predictor occasionally increases as the natural logarithm of lambda increases. Four predictors, including age of the forest owners, average annual income, contribution of off-farm income to annual income and contribution of crop income to annual income, have negative coefficients in the model.

According to the magnitude of the coefficient estimates, the predictors can be divided into three groups:

1. The first and most important group comprises the frequency of forest maintenance activities (Fmain), receiving technical support from forestry extension workers (Tsup), timber price (TprP) and contribution of crop income to annual income (Crpin) predictors as these predictors tend to have by far the largest coefficient estimates in the set.
2. The second group contains the total forestland plot (Tfp), family member available for planting trees (FmavP), total forestland area (Parcel), reason owning forestland for creating natural landscape (Ronlc), awareness about government subsidy grant for establishing forests (Govt), reason owning forestland for generating forestry income (Rogfi), contribution of forestry income to annual income (Frtin) and contribution of off-farm income to annual income (Othin) predictors.

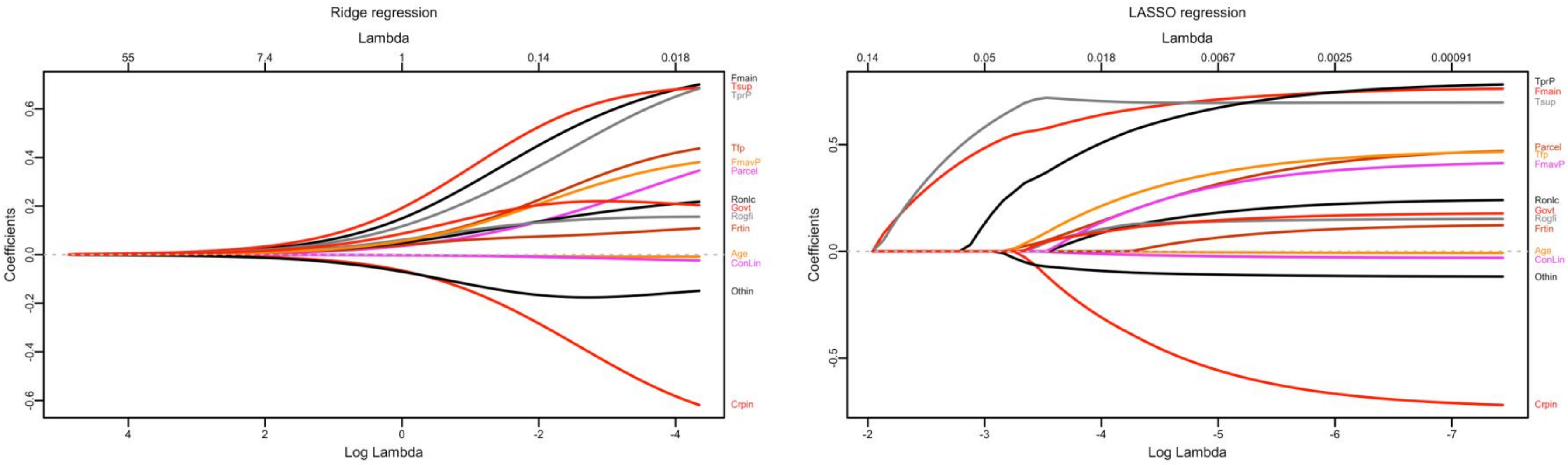
3. The last group contains the average annual income (ConLin) and age of the forest owners (Age) predictors since these predictors' coefficient estimates are effectively zero.

The right panel graphically presents the results of the LASSO regression. The importance of these predictors can be divided into three groups according their presence in the model as lambda is varied.

1. The first and most important group contains the timber price (TprP), frequency of forest maintenance activities (Fmain) and receiving technical support from forestry extension workers (Tsup) predictors, as these predictors' coefficient estimates are lastly shrunk to zero.
2. The second group contains total forestland area (Parcel), total forestland plot (Tfp), family member available for planting trees (FmavP), reason owning forestland for creating natural landscape (Ronlc), awareness about government subsidy grant for planting trees (Govt), reason owning forestland for generating forestry income (Rogfi), contribution of off-farm income to annual income (Othin) and contribution of crop income to annual income (Crpin) predictors. These predictors' coefficient estimates are simultaneously and/or quickly shrunk to zero at the same time as the natural logarithm of lambda increases from -4 to -3.
3. The last group contains age of the forest owners (Age), average annual income (ConLin) and contribution of forestry income to annual income (Frtin) predictor. These predictors' coefficients shrink to zero when natural logarithm of lambda is close to -4.

Figure 10 is a plot of a 10-fold cross-validation curve, with the red dotted lines and upper and lower standard deviation curves along the sequence of lambda. The top x-axis present number of predictors along the sequence of lambda. The top and lower panels present the cross-validation statistics for Ridge and LASSO, respectively. The two vertical dotted lines represent the value of lambda that gives the minimum misclassification error. The other dotted line represents the most regularised mode such that error is within one standard error from the minimum misclassification rate.

Figure 10 shows that both methods have a similar range of misclassification error, from 0.14 to 0.22. The minimum misclassification rate of both methods is about 16%. To produce the minimum misclassification rate both methods use all candidate predictors. However, the effects of penalty terms are clearly demonstrated when the misclassification error is greater than 0.18. When the misclassification error is above 0.18, to produce the same misclassification error, LASSO uses less predictors than Ridge.



Legend:	
Tfp:Total forestland plot	Ronlc: Reason owning forestland for creating natural landscape
Parcel: Total forest land area	Rogfi: Reason owning forestland for generating forestry income
ConLin: Average annual income	Fmain: Frequency of forest maintenance activities
Crpin: Contribution of crop income to annual income	Tsup: Receiving technical support from forestry extension workers
Frtin: Contribution of forestry income to annual income	FmavP: Family member available for planting trees
Othin: Contribution of off-farm income to annual income	Govt: Awareness about the government subsidy grant for establishing forests
Age: Age of the forest owners	TprP: Timber price

Figure 9: Coefficient regularised path of tree planting models

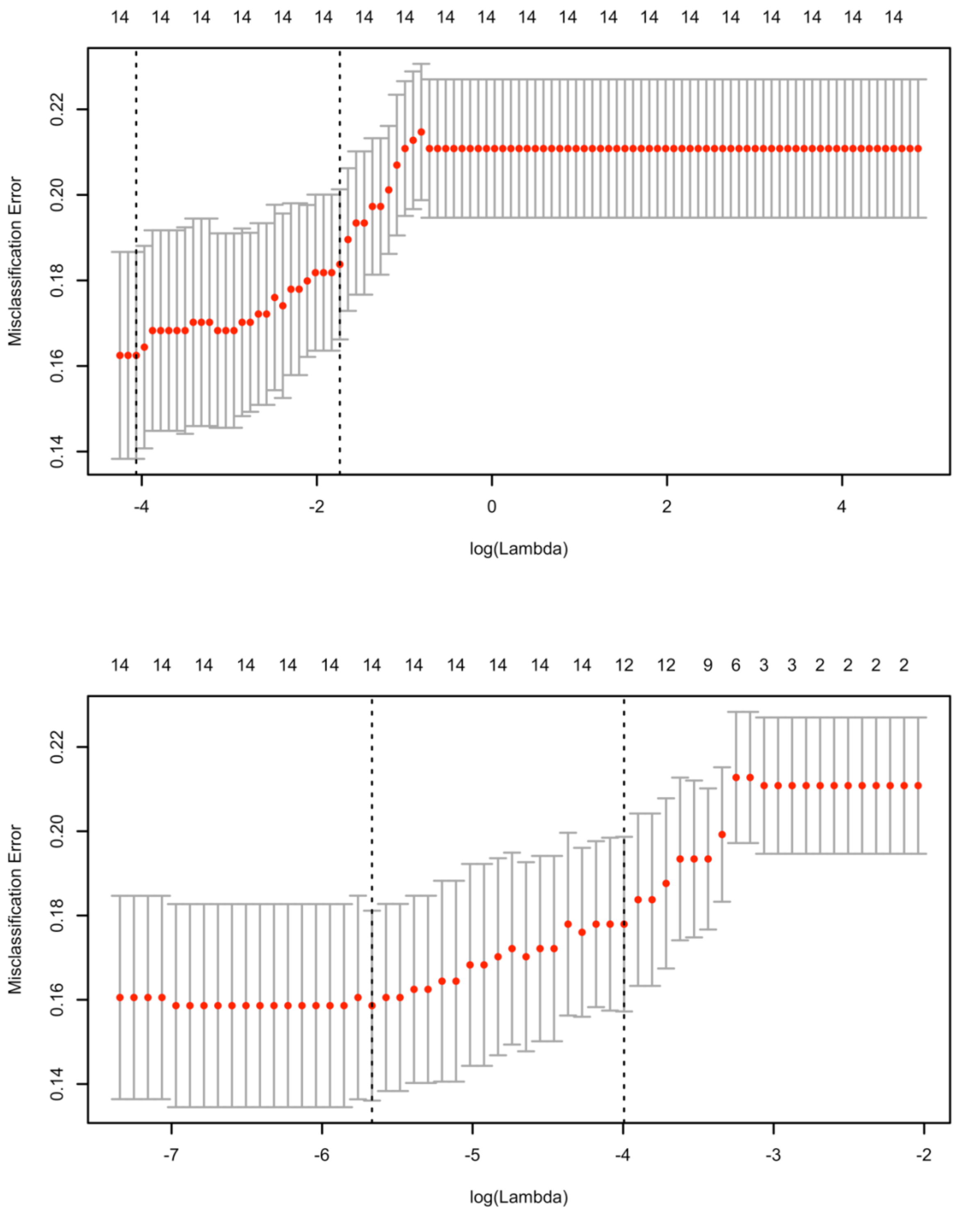


Figure 10: 10-fold cross-validation curve corresponding to a sequence of λ.

Top panel: Ridge regression. Lower panel: LASSO regression

4.5.5. Summary

The analysis results suggest that there is a statistical association between the forest owners' decisions in planting trees and the set of fourteen predictors. The result of the VIF analysis statistically indicates that the problem of high multicollinearity is not presented in this set of predictors.

The best subset selection, Ridge, and LASSO regression techniques were used to develop multiple predictor models and to rank the importance of predictors. The Ridge and LASSO regression results showcase a minimum misclassification rate of approximately 0.16 or 16%, using the 10-fold cross validation subsampling method. The optimal model of the best subset selection that was selected for use is an averaged model that was estimated in the multi-model inference framework. The misclassification rate of the averaged model is approximately18%. It can be concluded that the three methods produce similar misclassification rates.

Table 12 presents the ceteris paribus impacts of the identified variables using the three different regression approaches. For convenience, the predictors are presented in groups namely forestland management objectives, labour distribution, institutional factors and market conditions. The forestland management objective group contains the total forestland area, the total number of forestland plots, the reason owning forestland to create natural landscape and the reason owning forestland to generate forestry income. These variables are positively correlated with tree-planting decisions. A possible explanation of this is that forest owners are motivated by generating additional income or enjoy the aesthetic landscapes of their forests. Logically, owning more forestland area and plots and planting trees are the most direct way to achieve their objectives.

The labour distribution group contains the age of forest owners, the available family members for planting trees, the frequency of forest maintenance activities, average annual income, the contribution of forestry income to annual income, the contribution of crop income to annual income and the contribution of off-farm income to annual income predictors.

The presence and the positive/negative correlations of these variables to some extent illustrate two key fundamental principles of economics that are described by Mankiw (2018): "people face trade-offs" and the "cost of something is what you give up to get it". Given that personal resources are limited, if forest owners allocate their labour resources to crop and off-farm activities, they would have less time for forestry activities. Consequently, they might decide not to plant trees. In contrast, if the forest owners highly value cash income from forestry activities, obviously planting trees would be a better use of their labour resources.

In addition, tree planting activities are physically labour-intensive activities. This might explain why the older forest owners are not likely to plant trees. However, if the forest owners' families have able-bodied members, the forest owners' probability of planting trees would increase.

Surprisingly, the average annual income of the forest owners was found to have a negative impact on tree planting decisions. Note that tree planting requires substantial upfront investment such as buying seedlings and site preparation. Therefore, logically, larger annual incomes imply a larger capital base for tree planting. However, the results for this predictor are contrary to the aforementioned logic. A possible explanation is that the contribution of forestry income to the annual income of the forest owners with high annual incomes is less than other sources of income such as off-farm and crop income.

The ceteris paribus impact of the frequency of forest maintenance activities predictor on the decision also must be interpreted with caution because the finding indicates that a higher level of forest maintenance activities carried out by the forest owners correlates with a greater probability for tree planting. The impact of this predictor might be the result of tree planting decisions, as the forest owners are compelled to maintain their forests because they own them. Or it could be argued that the positive correlation was due to the forest owners considering forest maintenance activities a part of their forestland management objective to create natural landscapes.

The institutional factor group contains the awareness of the forest owners about the government subsidy grant for establishing forests and technical supports of forestry extension workers. These predictors are positively correlated with tree-planting decisions. These findings reflect the fourth principle of economics according to Mankiw (2018): "people respond to incentives". The financial support of the grant assists the forest owners by reducing the upfront investment in planting trees. Meanwhile, forestry extension workers who play a role as knowledge builders and motivators assist the forestry owners in understanding and efficiently using their forestland.

The last group contains only the timber price predictor. The findings indicate that the forest owners include the timber price in their decision. Since some forest owners seek income from forestry, knowledge of the timber price is one of indicators that they would plant trees.

Table 12: The optimal models for decision in tree planting

	Ridge*	**LASSO***	**Best subset selection**
	Estimates	*Estimates*	*Averaged estimates*
(Intercept)	-1.04	-1.12	0.89
Forestland management objectives			
Total forestland area	0.32	0.39	0.56
Total forestland plot	0.42	0.42	0.31
Reason owning forestland for creating natural landscape	0.21	0.21	0.16
Reason owning forestland for generating forestry income	0.16	0.15	0.09
Labour distribution			
Age of the forest owners	-0.01	-0.01	-0.001
Family member available for planting trees	0.37	0.36	0.26
Frequency of forest maintenance activities	0.68	0.74	0.78
Average annual income	-0.02	-0.03	-0.04
Contribution of forestry income to annual income	0.1	0.09	0.08
Contribution of crop income to annual income	-0.59	-0.64	-0.69
Contribution of off-farm income to annual income	-0.15	-0.11	-0.03
Institutional factors			
Awareness about government subsidy grant for establishing forests	0.21	0.16	0.04
Receiving technical support from forestry extension workers	0.68	0.69	0.8
Market conditions			
Timber price	0.66	0.73	0.9

*: *presented models with minimum misclassification rate using 10-fold cross-validation*

4.6. Tree Harvesting Decisions

4.6.1. Variable Identification

Table 13 presents a summary of harvesting logistic regressions with respect to the forest owners' land asset characteristics. The results show that the total forestland area of the forest owner was uncorrelated with tree-harvesting decisions.

The distance from the house of the forest owners to their closest forestland plot was negatively correlated with tree planting decision. The total number of forestland plots and the total crop land area were, on the other hand, positively correlated with tree-planting decisions. Therefore, with respect to the land asset characteristics of the forest owners, these predictors were chosen as candidate variables for developing the multiple-predictor model.

Table 13: Summary of tree harvesting logistic regression with respect to land asset

	CrpA (1)	Parcel (2)	Tfp (3)	Dist (4)
Total cropland area (CrpA)	5.92^{***} (1.75)			
Total forestland area (Parcel)		.09 (.15)		
Total forestland plots (Tfp)			1.02^{**} (.45)	
Distance from house to closest forestland plot (Dist)				$-.17^{**}$ (.08)
Constant	1.06^{***} (.27)	1.86^{***} (.18)	.73 (.51)	2.21^{***} (.20)
Observations	408	408	408	408
Log Likelihood	-148.61	-155.47	-151.39	-153.38
Akaike Inf. Crit.	301.23	314.93	306.79	310.76
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 14 shows a summary of logistic regressions for the decision to harvest trees with respect to the forest owner personal characteristics. The level of education was negatively correlated with tree-harvesting decisions. The remaining predictors were not correlated with harvesting decisions.

Table 14: Summary of tree harvesting logistic regression with respect to personal characteristics

	Age (1)	Gen (2)	edut (3)	Fco (4)	FcoM (5)	FcoF (6)	FcoW (7)	FcoPF (8)
Age	.01 (.01)							
Gender		.47 (.49)						
Education level			$-.53^{*}$ (.31)					
Total number of people in forest owner's family (Fco)				.10 (.13)				
Total number of males in forest owner's family (FcoM)					.05 (.16)			
Total number of females in forest owner's family (FcoF)						.12 (.17)		
Total number of people in the workforce in forest owner's family (FcoW)							.13 (.13)	
Total number of people particiapting in forestry activities in forest owner's family (FcoPF)								.22 (.13)
Constant	1.59^{***} (.59)	1.87^{***} (.16)	2.25^{***} (.25)	1.47^{**} (.63)	1.80^{***} (.45)	1.67^{***} (.39)	1.54^{***} (.42)	1.39^{***} (.35)
Observations	408	408	408	408	408	408	408	408
Log Likelihood	-155.49	-155.14	-154.17	-155.39	-155.61	-155.41	-155.17	-154.24
Akaike Inf. Crit.	314.97	314.29	312.35	314.78	315.22	314.82	314.33	312.49
Note:						$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$		

Table 15 presents a summary of logistic regressions for the decision to harvest trees with respect to forest owners’ structure of income. The contribution of crop income to annual income and the contribution of forestry income to annual income were positively correlated with harvesting decisions. The average annual income was, on the other hand, negatively correlated with harvesting decisions. These predictors were chosen as candidate predictors for developing multiple predictor models. The contribution of off-farm income to annual income was not correlated with harvesting decision.

Table 15: Summary of tree harvesting logistic regression with respect to structure of annual income

	ConLin (1)	Crpin (2)	Frtin (3)	Othin (4)
Average annual income (ConLin)	$-.02^{*}$ (.01)			
Contribution of crop income to annual income (Crpin)		$.62^{***}$ (.19)		
Contribution of foresry income to annual income (Frtin)			$.68^{***}$ (.21)	
Contribution of off-farm income to annual income (Othin)				-.07 (.25)
Constant	2.45^{***} (.33)	.08 (.55)	.01 (.59)	2.15^{**} (.83)
Observations	408	408	408	408
Log Likelihood	-153.81	-150.21	-150.47	-155.62
Akaike Inf. Crit.	311.63	304.42	304.95	315.24
Note:				$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table 16 summarises the results of different logistic regressions on six nominated variables representing the forest owners' management objectives. The reason owning forestland for future generation and receiving technical support from forestry extension workers were not correlated with harvesting decisions.

The reason owning forestland for creating natural landscape and the frequency of forest maintenance activities were negatively correlated with harvesting decisions. The reason owning forestland to generate forestry income and the reason for owning forestland as an investment were positively correlated with harvesting decisions. These predictors were chosen as candidate predictors for developing multiple predictor models.

Table 16: Summary of tree harvesting logistic regression with respect to forestland management objectives

	Decisions in Harvesting Trees					
	Roinv	Rokfg	Ronlc	Rogfi	Fmain	Tsup
	(1)	(2)	(3)	(4)	(5)	(6)
Reason owning forestland as an investment (Roinv)	$.41^{*}$ (.24)					
Reason owning forestland for future generation (Rokfg)		.23 (.25)				
Reason owning forestland for creating natural landscape (Ronlc)			$-.76^{***}$ (.19)			
Reason owning forestland for generating forestry income (Rogfi)				$.44^{**}$ (.22)		
The frequency of forest maintenance activities (Fmain)					$-.76^{***}$ (.27)	
Receiving technical supports from forestry extension workers (Tsup)						-.07 (.32)
Constant	.66 (.75)	1.19 (.83)	3.27^{***} (.39)	.61 (.66)	3.09^{***} (.47)	1.97^{***} (.27)
Observations	408	408	408	408	408	408
Log Likelihood	-154.22	-155.25	-147.48	-153.70	-151.51	-155.63
Akaike Inf. Crit.	312.45	314.50	298.96	311.39	307.02	315.26
Note:						*p<0.1; **p<0.05; ***p<0.01

Table 17 shows a summary of the different logistic regressions on seven categorical variables representing factors that the forest owners claimed to have effects on their harvesting decisions. Of these variables, the availability of family member for harvesting trees and the ability to access to the road for transporting harvesting products predictors were not correlated with harvesting decisions.

The remaining predictors including the cost of harvesting, the timber price, the government regulations on harvesting activities, the tree age and the keeping of forests as a type of financial saving that will be harvested when cash is needed were correlated with harvesting decision. These predictors were chosen as predictors for developing multiple predictor models.

Table 17: Summary of tree harvesting logistic regression with respect to other categorical factors

	FmavH (1)	Acrd (2)	Chvt (3)	TprH (4)	Greg (5)	Tage (6)	Csave (7)
The availability of family member for harvesting trees (FmavH)	-.05 (.31)						
The ability to access to the road for transporting harvesting products (Acrd)		.13 (.36)					
Cost of harvesting (Chvt)			1.36*** (.31)				
Timber price (TprH)				1.32*** (.37)			
Government regulations on harvesting activities (Greg)					.53* (.31)		
Tree age (Tage)						1.87*** (.32)	
Considering forest as a type of financial saving (Csave)							2.59*** (.36)
Constant	1.94*** (.19)	1.90*** (.17)	1.15*** (.21)	.80** (.33)	1.72*** (.18)	.82*** (.21)	-.14 (.31)
Observations	408	408	408	408	408	408	408
Log Likelihood	-155.64	-155.60	-145.64	-150.15	-154.17	-137.29	-130.96
Akaike Inf. Crit.	315.29	315.19	295.29	304.31	312.35	278.57	265.92

Note: *p<0.1; **p<0.05; ***p<0.01

4.6.2. Variance Inflation Factor (VIF)

The results of the multicollinearity analysis are presented in Tables 18 and 19. The results suggest that there is no existence of high multicollinearity among candidate predictors. The minimum correlation was 0.0007, which occurs between tree age and the contribution of forestry income to annual income. The maximum correlation was 0.66, which occurs between reason owning forestland for generating forestry income and reason owning forestland as an investment. All candidate predictors were used for building the final logistic model.

Table 18: VIF values for the harvesting-decision candidate predictors

No.	Predictors	VIF values
1	Total forestland plot (Tfp)	1.15
2	Total cropland area (CrpA)	1.92
3	Average annual income (ConLin)	2.07
4	Contribution of crop income to annual income (Crpin)	2.21
5	Contribution of forestry income to annual income (Frtin)	2.08
6	Education level (edut)	1.16
7	Distance from forest owners’ house to closest forestland plot (Dist)	1.24
8	Reason owning forestland as an investment (Roinv)	2.38
9	Reason owning forestland for creating natural landscape (Ronlc)	1.46
10	Reason owning forestland for generating forestry income (Rogfi)	2.11
11	Frequency of forest maintenance activities (Fmain)	1.16
12	Cost of harvesting (Chvt)	1.62
13	Timber price (TprH)	1.22
14	Government regulations on harvesting activities (Rreg)	1.39
15	Tree age (Tage)	1.84
16	Considering forests as a type of financial saving (Csave)	1.27

Table 19: Correlation matrix for harvesting-decision candidate predictors

	Tfp	*CrpA*	*ConLin*	*Crpin*	*Frtin*	*edut*	*Dist*	*Roinv*	*Ronlc*	*Rogfi*	*Fmain*	*Chvt*	*TprH*	*Greg*	*Tage*	*Csave*
Tfp																
CrpA	-0.11^{*}															
ConLin	0.22^{***}	-0.46^{***}														
Crpin	0.01	0.59^{***}	-0.42^{***}													
Frtin	0.04	0.45^{***}	-0.21^{***}	0.55^{***}												
edut	-0.11^{*}	-0.01	0.12^{*}	-0.16^{**}	-0.01											
Dist	-0.10^{*}	-0.05	0.29^{***}	-0.12^{*}	-0.10^{*}	0.15^{**}										
Roinv	-0.05	0.41^{***}	-0.30^{***}	0.54^{***}	0.60^{***}	-0.06	-0.09									
Ronlc	-0.23^{***}	0.23^{***}	-0.17^{***}	0.05	0.18^{***}	0.20^{***}	0.14^{**}	0.31^{***}								
Rogfi	-0.00	0.37^{***}	-0.26^{***}	0.49^{***}	0.60^{***}	-0.03	-0.08	0.66^{***}	0.19^{***}							
Fmain	-0.00	-0.22^{***}	0.26^{***}	-0.21^{***}	-0.13^{**}	0.12^{*}	0.21^{***}	-0.18^{***}	-0.03	-0.05						
Chvt	0.19^{***}	-0.07	0.25^{***}	-0.02	0.09	-0.07	-0.07	0.01	-0.22^{***}	0.05	0.01					
TprH	0.06	-0.01	0.04	0.07	0.08	-0.12^{*}	0.03	0.09	-0.12^{*}	0.09	-0.06	0.29^{***}				
Greg	-0.12^{*}	0.22^{***}	-0.40^{***}	0.10^{*}	0.01	0.03	-0.10^{*}	0.10^{*}	0.14^{**}	0.14^{**}	-0.01	0.08	0.02			
Tage	0.19^{***}	-0.09	0.19^{***}	-0.01	-0.00	-0.19^{***}	-0.07	-0.10^{*}	-0.32^{***}	-0.01	0.02	0.56^{***}	0.36^{***}	0.16^{**}		
Csave	0.09	0.20^{***}	-0.16^{***}	0.19^{***}	0.12^{*}	-0.16^{**}	-0.13^{**}	0.05	-0.22^{***}	0.07	-0.12^{*}	0.18^{***}	0.20^{***}	0.09	0.30^{***}	

Computed correlation used pearson-method with listwise-deletion.

4.6.3. Best Subset Method

The set of sixteen candidate predictors was used to develop the multiple predictor models by applying Algorithm 2. The top part of Table 20 presents general information about the set of 100 candidate models. Figure 11 graphically presents the result of the best subset selection procedure. The best and the worst model's AIC are 242 and 244 respectively. All candidate models are within two units of each other in terms of AIC. It suggests that any model in the set can be chosen as a model for approximating the harvest decision. Therefore, multi-model inference framework was used to arrive the final multiple predictor model. The result of the multi-model inference is presented in lower part of Table 20.

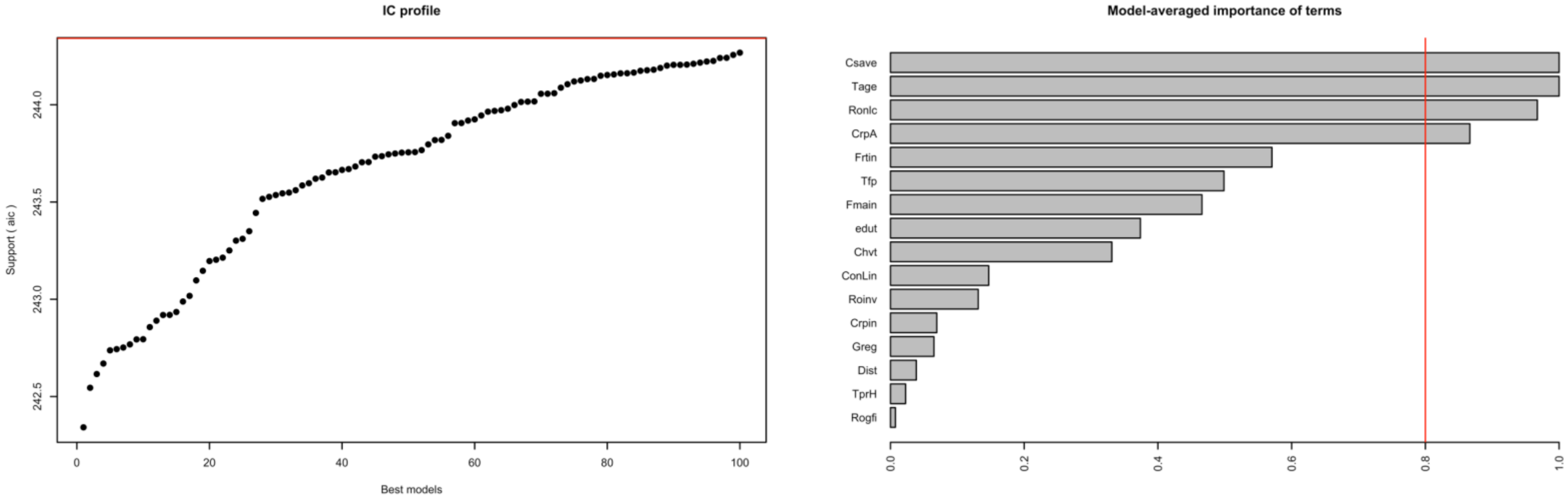
Note that five out of sixteen predictors have negative coefficients in the averaged model. The averaged model correctly predicts 365 out 408 cases, or equivalently, a misclassification rate of 12%. The average importance of each model's terms can be divided into three groups.

1. The most important group contains the total cropland area, reason owning forestland for creating natural landscape, considering forests as a type of financial saving and tree age predictors. These predictors have level of importance above 0.8.
2. The second important group contains the contribution of forestry income to annual income, total forestland plot, frequency of forest maintenance activities, education level and cost of harvesting predictor. These predictors have level of importance from 0.3 to 0.5.
3. The last group contains the average annual income, reason owning forestland as an investment, government regulations on harvesting activities, contribution of crop income to annual income, distance from the forest owners' house to their closest forestland plot, reason owning forestland for generating forestry income and timber price predictors. These predictors have level of importance below 0.3.

Table 20: Summary of the best subset selection procedure for harvesting decisions

glmulti.analysis
Method: h / Fitting: glm / IC used: aic
Level: 1 / Marginality: FALSE
From 100 models:
Best IC: 242.34
Worst IC: 244.27
100 models within 2 AIC units.

predictors	***Averaged Estimates***	***Nb models***	***Importance***
(Intercept)	-0.809	100	
Tree age (Tage)	1.301	100	1
Considering forests as a type of financial saving (Csave)	1.699	100	1
Reason owning forestland for creating natural landscape (Ronlc)	-0.51	96	0.967
Total cropland area (CrpA)	3.295	86	0.867
Contribution of forestry income to annual income (Frtin)	0.249	57	0.571
Total forestland plot (Tfp)	0.3	51	0.499
Frequency of forest maintenance activities (Fmain)	-0.209	45	0.466
Education level (edut)	0.196	39	0.374
Cost of harvesting (Chvt)	0.186	33	0.331
Average annual income (ConLin)	-0.003	18	0.147
Reason owning forestland as an investment (Roinv)	0.044	14	0.131
Contribution of crop income to annual income (Crpin)	-0.018	8	0.069
Government regulations on harvesting activities (Greg)	0.02	8	0.065
Distance from forest owners' house to their closest forestland plot (Dist)	-0.003	5	0.039
Timber price	0.007	3	0.023
Reason owning forestland for generating forestry income (Rogfi)	0.001	1	0.007



Legend:	
Tfp: Total forestland plot	Ronlc: Reason owning forestland for creating natural landscape
CrpA: Total cropland area	Rogfi: Reason owning forestland for generating forestry income
ConLin: Average annual income	Fmain: Frequency of forest maintenance activities
Crpin: Contribution of crop income to annual income	Chvt: Cost of harvesting
Frtin: Contribution of forestry income to annual income	TprH: Timber price
edut: Education level	Greg: Government regulations on harvesting activities
Dist: Distance from forest owners' house to closest forestland plot	Tage: Tree age
Roinv: Reason owning forestland as an investment	Csave: Considering forests as a type of financial saving

Figure 11: Graphical results of the best subset selection for decisions in harvesting trees.

Top panel: AIC profile. Lower panel: Estimated importance of predictors.

4.6.4. Shrinkage Regression

The results of shrinkage regression are presented in Figure 12. The left panel shows that the Ridge coefficient estimates tend to decrease in aggregate as the natural logarithm of lambda increases. Some of the coefficients' estimates of the timber price predictor (TprH) increase as the natural logarithm of lambda increases. It can also be observed that when the natural logarithm lambda is approximately greater than 1, the coefficients' estimates of the predictors are effectively zero.

Interestingly, the contribution of crop income to annual income predictor (Crpin) initially has a negative coefficient in the model. However, when the natural logarithm of lambda increases approximately above -2 the sign of this predictor has a positive coefficient in the model. According to the magnitude of the coefficients' estimates, the predictors can be divided into three groups.

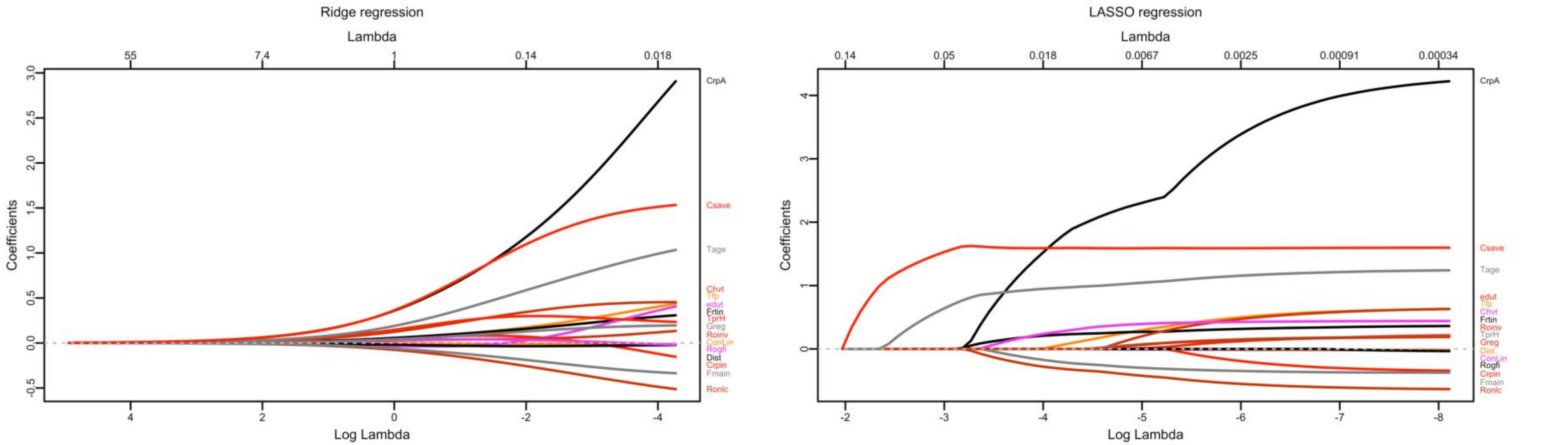
1. The first and most important group contains the total cropland area (CrpA), considering forests as a type of financial saving (Csave), tree age (Tage) and reason owning forestland for creating natural landscape (Ronlc) predictors since these predictors tend to have the largest coefficient estimates in the set.
2. The second group contains the cost of harvesting (Chvt), total forestland plot (Tfp), education level (edut), contribution of forestry income to annual income (Frtin), timber price (TprH), government regulation on harvesting activities (Greg) and frequency of forest maintenance activities (Fmain) predictors.
3. The last group contains the reason owning forestland as an investment (Roinv), reason owning forestland for generating forestry income (Rogfi), average annual income (ConLin), distance from the forest owners' house to their closest forestland plot (Dist) and contribution of crop income to annual income (Crpin) predictors since these predictors' coefficient estimates are basically zero.

In the right panel, showing a graphical result of LASSO, note that the contribution of crop income to annual income (Crpin) predictor did not change its sign in the model as it did in Ridge regression. The importance of predictors can be classified into three groups.

1. The most importance group contains the considering forests as a type of financial saving (Csave) and tree age (Tage) predictor since these predictors are the last ones to be eliminated from the model.

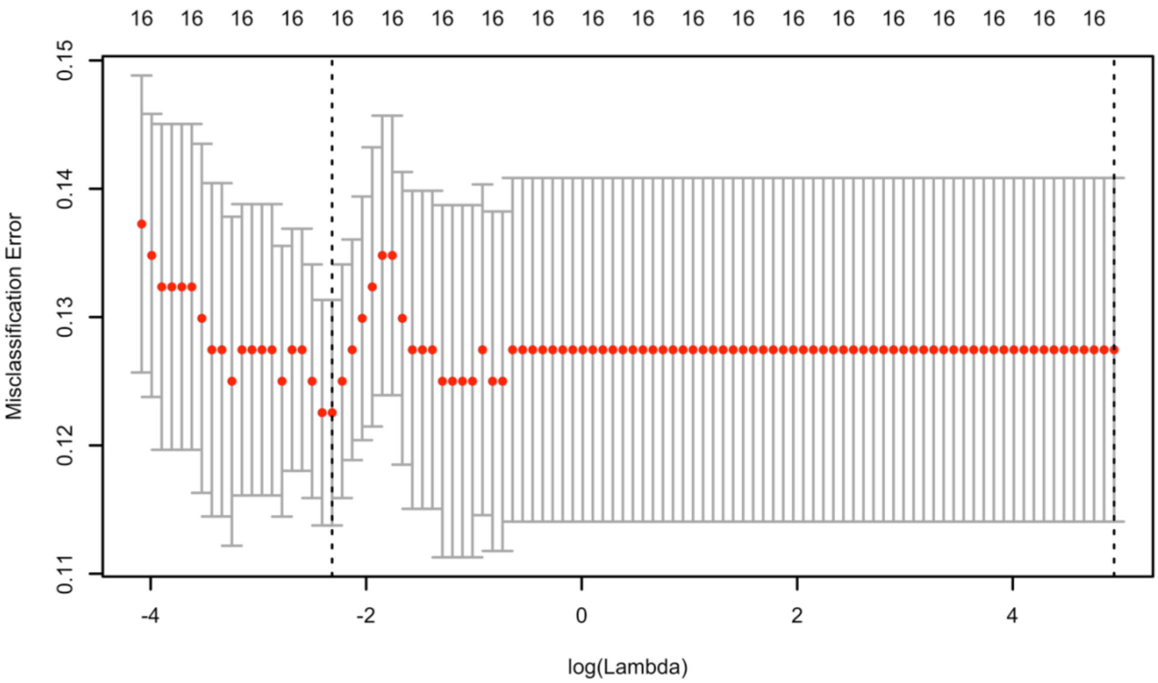
2. The second group contains the total forestland plot (Tfp), cost of harvesting (Chvt), contribution of forestry income to annual income (Frtin), frequency of forest maintenance activities (Fmain) and reason owing forestland for creating natural landscape (Ronlc) predictor. The predictors in this group are simultaneously or quickly eliminated from the model as the natural logarithm of lambda increases from -4 to -3.
3. The last group contains the government regulations on harvesting activities (Greg), timber price (TprH), reason owning forestland as an investment (Roinv), distance from the forest owner to their closest forestland plot(Dist), contribution of crop income to annual income (Crpin) and average annual income (ConLin) predictors since these predictors' coefficient estimates are eliminated from the model before the natural logarithm of lambda reaches -4.

Figure 13 is a plot of the 10-fold cross-validation curve along the sequence of Lambda with respect to Ridge and LASSO regression. As can be seen from the plot, the Ridge and LASSO regression misclassification varies from 0.11 to 0.15. When the natural logarithm of lambda is greater -4, LASSO regression uses fewer predictors than Ridge regression in order to produce a similar classification rate. The minimum misclassification rate of both methods is approximately 12%.



Legend:	
Tfp: Total forestland plot	Ronlc: Reason owning forestland for creating natural landscape
CrpA: Total cropland area	Rogfi: Reason owning forestland for generating forestry income
ConLin: Average annual income	Fmain: Frequency of forest maintenance activities
Crpin: Contribution of crop income to annual income	Chvt: Cost of harvesting
Frtin: Contribution of forestry income to annual income	TprH: Timber price
edut: Education level	Greg: Government regulations on harvesting activities
Dist: Distance from forest owners' house to closest forestland plot	Tage: Tree age
Roinv: Reason owning forestland as an investment	Csave: Considering forests as a type of financial saving

Figure 12: Coefficient-regularised path of tree harvesting models



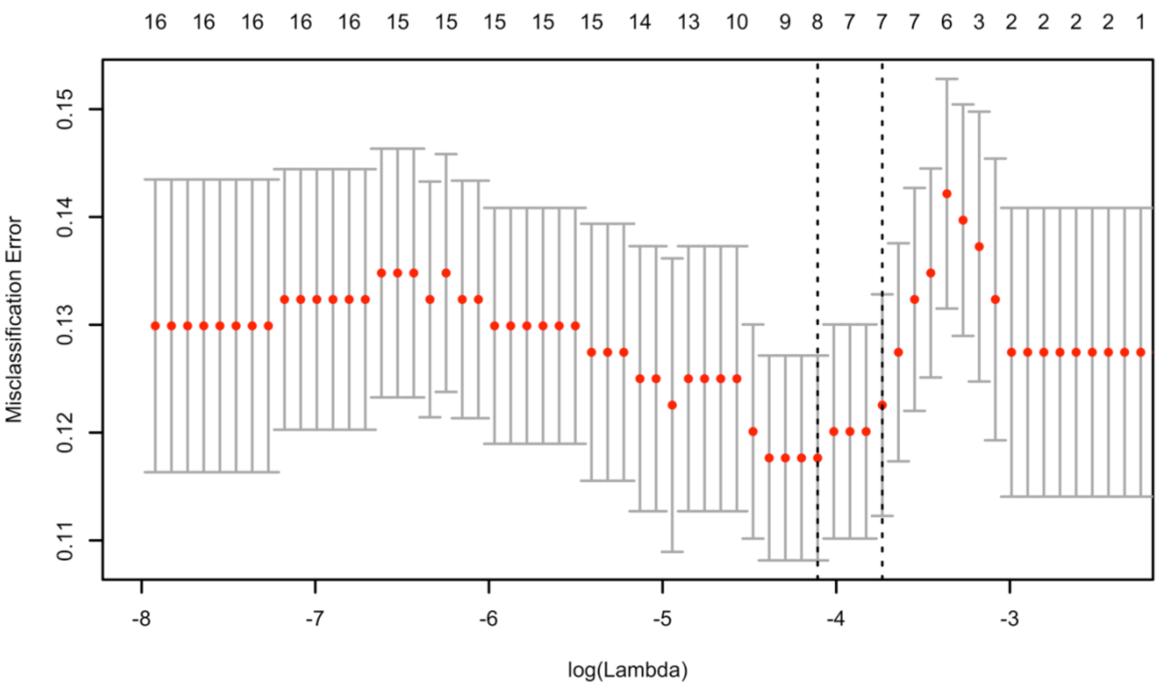


Figure 13: 10-fold cross-validation curve corresponding to a sequence of λ.

Top panel: Ridge regression. Lower panel: LASSO regression

4.6.5. Summary

The results of the tree-harvesting decision analysis show that there is a statistical association between the tree-harvesting decisions and the set of sixteen predictors. The results of the VIF analysis indicate that the problem of multi-collinearity between these variables is not present.

The Ridge and LASSO approaches resulted in a minimum misclassification rate of approximately 12% when using the 10-fold cross validation method. This minimum misclassification rate was obtained when Ridge regression employed all available variables. In contrast, the LASSO method employed only eight variables when the minimum rate was obtained. The model parameters of the best subset selection are produced by averaged estimates from a set of 100 candidate models. When the multi-model inference method was used, a misclassification rate of approximately 11% was obtained.

Table 21 presents the ceteris paribus impacts of the identified predictors on the harvesting decision-making with respect to the three regression methods. For convenience, the predictors are presented in groups namely: forest management objectives, labour distribution, market conditions, institutional factors, tree biological factors and personal financial conditions.

The forestland management objective group contains the total forestland plots, the distance from the forest owner's house to their closest forest plot, the reason for owning forestland as an investment, the reason for owning forestland for creating natural landscapes and the reason for owning forestland for generating forestry income predictors.

The total forestland plots, reason owning forestland as an investment and reason owning forestland for generating forestry income predictors positively correlated with harvesting decisions. The distance from the forest owners' house to their closest forestland plot and reason owning forestland for creating natural landscape were, on the other hand, negatively correlated with harvesting decision.

Table 21: The optimal models for decision in harvesting trees

	Ridge*	LASSO*	Best subset selection
	Estimates	*Estimates*	*Averaged estimates*
(Intercept)	-3.16	-3.06	-0.81
Forestland management objectives			
Total forestland plot	0.08	0.04	0.3
Distance from forest owners' house to their closest forestland plot	-0.03		-0.003
Reason owning forestland as an investment	0.05		0.04
Reason owning forestland for creating natural landscape	-0.29	-0.29	-0.51
Reason owning forestland for generating forestry income	0.03		0.001
Labour distribution			
Total cropland area	1.37	1.7	3.3
Average annual income	-0.006		-0.003
Contribution of crop income to annual income	0.06		-0.02
Contribution of forestry income to annual income	0.18	0.21	0.25
Frequency of forest maintenance activities	-0.21	-0.19	-0.21
Education level	0.07		0.19
Institutional factors			
Government regulations on harvesting activities	0.14		0.02
Market conditions			
Cost of harvesting	0.37	0.26	0.19
Timber price	0.29		0.01
Tree biological factor			
Tree age	0.65	0.96	1.3
Personal financial conditions			
Considering forests as a type of financial saving	1.2	1.59	1.7

**: presented models with minimum misclassification rate using 10-fold cross-validation*

The presence and sign of these predictors are logically supported. The forest owners who own the forestland as an investment or for generating income from forestry activities would harvest their forests to achieve their objectives. Similarly, the forest owners who want to enjoy their natural landscapes would obviously not want to harvest their forests.

The impact of distance from forest owners' house to their closest forestland plot on the decision needs to be interpreted with caution since the current structure of the data does not provide a reasonable explanation why a larger distance from the forest owner's house to the closest forest plots results in the reduction of probability to harvest the forests. The aforementioned uncertainty can be avoided if the LASSO regression is used. LASSO regression did not employ this predictor in the model but still produced a similar misclassification rate.

The labour distribution group contains the total crop land area, the average annual income, the importance of crop income to annual income, the contribution of forestry income to annual income, the frequency of forest maintenance activities, and the level of education of the forest owners. It can be observed that the magnitude of the ceteris paribus impacts of the average annual income, contribution of crop income and education level on the probability of harvesting trees is small. With respect to Ridge regression, if the value of the average annual income predictor increases by a unit, the harvest probability reduces by 0.6%. If contribution of crop income to annual income or education level predictor increases by a unit, the harvest probability increases by 6% or 7% respectively. A similar pattern is found in the averaged model.

As expected, these variables are eliminated from the LASSO model. The impact of the contribution of forestry income to annual income predictor is also logically supported: If the forest owners depend strongly on the income generated from forestry activities, obviously they will have a strong motivation to harvest their forests.

The remaining variables in this group, the total cropland area and frequency of forest maintenance activities predictors, have a greater influence on harvesting decisions than the other. However, the impact of these predictors must be interpreted with caution because with the current data structure, it is challenging to draw inferences on why an increase in crop area results increased harvesting probability.

The frequency of forest maintenance activities predictor is negatively correlated with tree-harvesting behaviour. This may be because of the fact that more than 50% of plots were planted with acacia. As mentioned in section 3.4.3, with respect to acacia plantation very little work is required since year 3 to harvesting stage. Therefore, the negative correlation is understandable.

The market conditions group comprises the cost of harvesting and the timber price predictor. Interestingly, the LASSO regression only employs the cost of harvesting in the model. The behaviour of timber price in the models, in combination with the third principle of economics (Mankiw, 2018) that “rational people think at the margin” suggests that the forest owners are more likely concerned about profit margins from harvesting forests instead of overall return from harvesting.

The group of institutional, biological and personal financial factors comprises the governmental regulations on tree harvesting activities, tree age and keeping forests as a type of financial saving predictor. Since forestry trees have a commercial life cycle, taking tree age into consideration is an expected result. A possible explanation of the presence of keeping tree as a type of financial saving predictor in the models is that the average annual income of the forest owners is approximately 21 million VND/year. This is a relatively low amount, suggesting that the level of annual saving is also low. In addition, the harvesting of forests provides them with a substantial one-time payment. Therefore, considering the forests as a source of financial savings is understandable.

4.7. Conclusions

This study has identified variables affecting decisions in tree planting and harvesting. For convenience, these variables were classified into groups: forestland management objectives, labour distribution, market conditions and institutional factors. The harvesting decisions are also affected by the personal financial condition factor and tree biological factors.

The predictors in the group of forestland management objectives and labour distribution play the most important role in the models. This is because there is at least one predictor that belongs to these groups that is present in both harvesting and planting models with respect to three regression methods. The forestland management objectives express the underlying reasons for owning forestland as well as the size of the forestland. For example, the forest owners who own forestland for generating income from forests and for enjoying natural landscapes tend to plant trees.

To understand the role of labour distribution in the forestry industry, it is necessary to know that the main characteristics of the forestry industry are that it has a long-term commercial life cycle and is highly labour intensive. Labour intensity can be seen from two perspectives. Firstly, the labour resource is required in the first two or three years to establish the forests. Secondly, the labour resource is required to protect the forests from various risks such as bushfire, theft, pests and disease. As noted by Byron (2001b) tree-farmers will not invest in a crop unless they are reasonably confident that their investment will survive to marketable maturity. Furthermore, the primary return on investment is normally collected as a one-time payment.

Because of these reasons, older forest owners (who tend to be in poorer physical condition) might tend to not plant trees. However, the presence of physically capable family members also affects their tree-planting decisions. In addition, the forest owners who are dependent on off-farm or crop income are less likely to plant trees since their labour resources are allocated to off-farm and crop activities. Thus, they do not have time to take care and protect their forests.

Institutional factors play significant roles in increasing the probability of planting trees but not the harvesting of trees. The market conditions group contains the timber price and the cost of harvesting variables. Interestingly, the timber price predictor has a less important role in explaining harvesting decisions than the cost of harvesting. This implies that the forest owners think more about profit margins than overall revenues. Besides this, harvesting decisions are also affected by tree age and the financial situation of the forest owners.

The analysis results provide evidence that the presence of these variables in the models, to some extent, reflects fundamental economic principles such as trade-offs, opportunity costs, responding to the incentives and maximising margin benefits. The forest owners make choices in accordance with these principles in order to maximise their well-being given that resources are limited. Understanding these strategies may offer policymakers a great tool for managing forests in the future.

Three regression techniques were used to obtain the multiple predictor models and to quantify the importance of individual predictors in the model. All three techniques yielded models that have effectively equivalent predictive ability. These methods are a great replacement for traditional stepwise selection because they not only perform model selection but also provide a means of quantifying the importance of predictors. An additional advantage of these methods is their ability to address a critical challenge in statistical model selection, which is to balance between bias and variance, since additional penalties are added to the original regression estimator.

The best subset selection method in combination with the multi-model inference framework and the Ridge regression approach yielded global models that contains all of the identified predictors. In contrast, LASSO yielded sparse models while also performing variable selection. Therefore, the LASSO method seems to be superior for the purposes of this research.

4.8. Limitations

The findings of this study may be somewhat limited in some aspects. Firstly, the models do not take geographical effects into account. As mentioned earlier, the information was collected from 517 people who were from different locations in the province. The modelling method in this study assumes that the behaviour of the forest owners does not vary by location. Future research could focus on modelling the variation of forest owners' behaviour by location and geography, which can then be compared to the aggregated results of this study.

The second limitation is about the level of detail of the final models. The final predictive models use categorical predictors such as the tree age, cost of harvesting and timber price. Therefore, the final models can only indicate that these predictors are included in the owners' decisions, but not specifically how they affect the decisions. This means that it is unclear at what tree age and timber price the owners will decide whether to harvest or not to harvest trees. Future research into these variables and incorporating them explicitly in the models would be useful.

The last limitation is about the use of the "tree planting behaviour" term. In this research, the term "tree planting" is used to refer to both reforestation and afforestation. From a forestry management point of view, afforestation behaviour plays a more important role than reforestation behaviour because afforestation behaviour is a better indicator of the expansion of new forests as compared to reforestation behaviour. Therefore, future research should separate and focus on afforestation behaviour rather than tree planting behaviour in general.

CHAPTER 5: MODELING THE AFFORESTATION AND HARVESTING INTENSITY OF NIPF OWNERS USING LINEAR MIXED – EFFECTS MODELS

5.1. Introduction

The previous chapter of this thesis identified the factors influencing the tree planting and harvesting decisions of the forest owners in Thai Nguyen province. However, the first study had revealed issues with the developed models that have yet to be addressed: 1) The models could not quantify the rate of change in forested areas, and 2) The models generalise tree planting decisions by grouping afforestation and reforestation into a single decision.

From a forest management point of view, afforestation plays a more important role than reforestation since afforestation is a better indicator of the expansion of new forests as compared to reforestation. The second study therefore attempts to model the afforestation and harvest intensity of the forest owners with the purpose of addressing the limitations of the models developed in the previous study.

This chapter presents the design and results of the second study. The chapter comprises five sections. Section 5.2 describes the sampling procedure and data collection. Section 5.3 explains the model fit. Section 5.4 presents the collected interview forms. Section 5.5 and 5.6 presents the results afforestation and harvesting model, respectively. Section 5.7 is conclusion.

5.2. Sampling Procedure and Data Collection

The second survey used the dataset from the Provincial Forest Inventory Project (PFIP) that was conducted in 2016 for purposes of developing the sampling procedure. The dataset contains information about the forest owners in the province such as total forestland area, types of forests and ownership type. Note that three districts of the province were purposely excluded from the dataset. Vo Nhai District was excluded as a large proportion of its forests and forestland belongs to the Than Sa - Phuong Hoang Nature Reserve which is under the management of People's Committee of Thai Nguyen. The city of Thai Nguyen itself and the Song Cong District were excluded due to their small proportion of forests and forestland area. The dataset therefore only comprises information on six districts in Thai Nguyen province.

The sampling procedure is affected by the nature of the model, the decision-making time of the forest owners and the geographical representation of the participants in the pool. As mentioned, the main objective of the study is to model the afforestation and harvesting intensity of the forest owners in Thai Nguyen province. Hence, the sample should include participants who can provide information that can be incorporated into the afforestation and harvesting models. Unfortunately, the dataset does not directly provide an indicator for determining the harvesting activities. The dataset does, however, provide the year in which the forest owner reforested and/or afforested their forestland. Therefore, we assumed that the forest owners harvested and replanted in the same year in order to derive an indicator for the harvesting activities. It was therefore assumed that the reforestation year and the year of harvest are equivalent.

The afforestation and harvest decision-making times are important inputs for the models, which require the use of historical information about the forest owners. The dataset was completed in 2016, which means that the most recent information about decision-making times was recorded that year. Therefore, to reduce selection bias with respect to the decision time points, the study only includes people who made afforestation and harvesting decision since 2014, assuming that the information on decisions before 2014 was more likely to be inaccurate due to the decreasing fidelity of the owners' recollections of their activities with time. Lastly, the participant's locations should represent the various geographical regions in the province. In summary, the sample should include the forest owners who

1. Have afforested their forestland since 2014.
2. Have available forestland for afforestation or never planted trees.
3. Have harvested their forests since 2014.
4. Have available forests for harvesting but did not harvest or never harvested trees
5. Represent the different geographical regions of the province.

If the participants are selected as a sampling unit according to the above criteria, the cost of data collection is increased due to spreading out the forest owners in the province. Therefore, instead of using the research subject as a sampling unit, the research used a so-called Primary Sampling Unit as described by Lavrakas (2008) for sampling procedure. This involves grouping the participants into aggregates according to their locations. The aggregated locations were used as the primary sampling unit.

The final six chosen locations were places that met all of the aforementioned criteria and that had the maximum number of potential interviewees: Bao Cuong, Boc Nhieu, Hop Tien, Minh Lap, Tien Hoi and Yen Do. These locations are indicated in Figure 14.

This study used interviews as the main tool to collect data, where the variables identified in the first study were used to construct the questions for the interview questionnaire. These variables are detailed in Appendix 5. The procedure of designing the questionnaire was similar to the first study that was described in previous chapter. The second survey questionnaire was subsequently sent to Human Ethics Committee of the University of Canterbury for approval. The research was granted approval on 9 August 2017 by the Chair of Human Ethics Committee of the University of Canterbury.

The final survey instrument was composed of four main sections. The first section collects information about the forest owner's personal characteristics such as age, gender, family structure and annual income. The second section contains questions about forests and forestland assets, and the management objectives of the forest owners. The third and fourth sections were designed to collect information about afforestation and tree harvesting respectively. Most of the questions in the survey were structured questions, with a few open-ended questions that were expected to be used as in-depth explanation for the models.

Three research assistants were employed to support the interviewing. These assistants were trained for a week in order to ensure that they understood the nature of the research, knew the structure of the questionnaire and possessed the required interview skills. This training was necessary as it further ensured the reliability of the results collected by the research assistants, as well as facilitated a smooth interview process.

Before conducting the interviews, the researcher arranged meetings with the local authorities and heads of villages of the locations where targeted participants reside. In these meetings, the researcher asked for consent to conduct the research in the regions and scheduled the interviews. In the end, the surveys were conducted in five out of the six initially selected locations, as consent was not given to carry out the research in one of the locations (Tien Hoi – Dai Tu District).

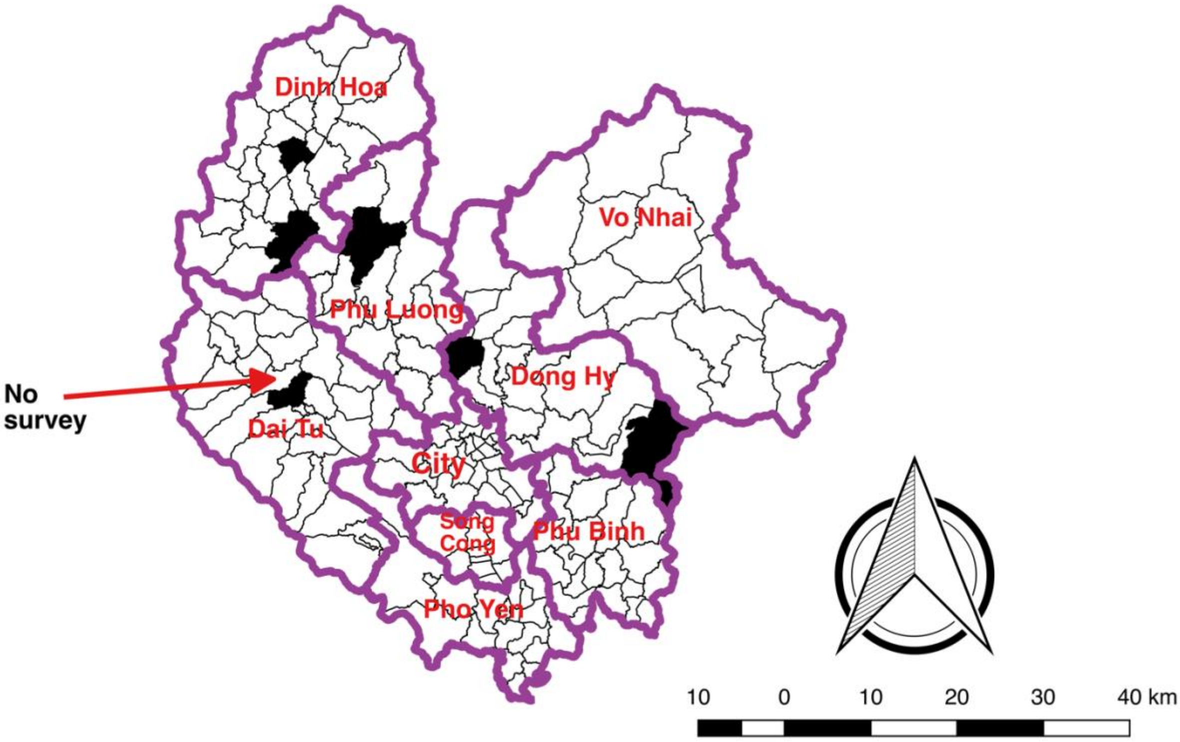


Figure 14: The location of the surveys

5.3. Model Construct

The purpose of this research is to develop an afforestation and harvesting intensity model. The structure of this model depends on several factors, including the nature of the dependent variable, the sampling method and the structure of captured data. The model's dependent variables are afforestation and harvesting intensity. The intensity is the proportion of afforested or harvested area versus the total available forestland or forest area for afforestation or harvest respectively. The independent variables continuously vary on a scale from zero to one. Thus, the model should be a regression-type instead of a classification-type as it was in the first survey.

The research uses participant's locations as primary sampling unit for sampling procedure. According to Lavrakas (2008), this type of sampling method introduces a hierarchical structure to the sample. In addition, the research includes information about the decisions of the forest owners since 2014. As a result, the collected information might be repeated more than one time in the dataset. This data structure is known as panel data or longitudinal data.

The model should have the capacity to deal with one of the common challenging issues in panel data: "incomplete" or panel drop out data, which is a result of randomly missing, unavailable or unobtainable records. For example, a forest owner has forestland plots and he/she afforested all plots in 2014. In this case, no further information on the forest owner was recorded in later years. Consequently, the information of this owner in dataset will be unbalanced in comparison with other owners who potentially afforested in later years.

According to Hox and Roberts (2011) and Gelman and Hill (2007), the linear mixed-effects model or the hierarchical/multi-level model construct can address all of the aforementioned challenges. In this research, a linear mixed-effect model construct was selected for data analysis. The models use the locations of the forest owners and the repetition of the forest owners in the dataset as random effects.

The fitting routine is conducted by the lme4 R package (Douglas Bates 2015). A maximum likelihood estimator was used to estimate model’s parameters. The p-value and model summary were generated by the SjPlot R-package (Lüdecke 2018). P-value was computed via a Wald-statistics approximation.

5.4. The Collected Interview Forms

The physical interview forms were transformed into electronic ones and subsequently transferred to an SQL database. Before that, a number of collected interview forms were removed from the sample as they were incomplete or there was mismatching between sections in the form.

A total of 318 usable interview forms from the five different locations were collected, with 179 male and 139 female participants. The distribution of participants in each of the five locations is presented in Figure 15. Of the 318 participants in the sample, 148 had completed high school or higher education, with the other participants possessing an intermediate school degree or a lower level of education.

Table 22 summaries the key statistical information of participants in the survey. The average age and the number of years each participant spent as a forest owner were 46 and 13 years respectively. Each participant owns an average of two plots of forestland with a total forestland area of 1.9 hectares. In general, the participants owned less cropland than forestland, with each participant owning an average of 0.2 hectares of cropland and some participants not owning any cropland at all.

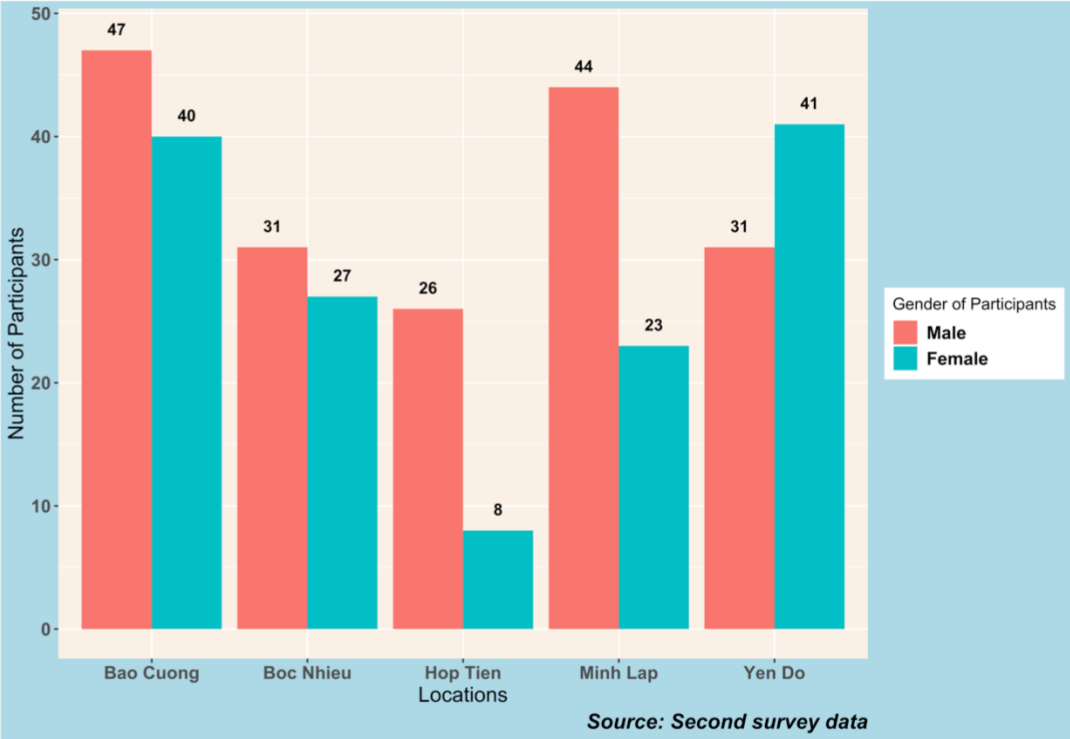


Figure 15: Distribution of the interviewees in the second survey

Table 22: Descriptive statistics of the second survey's participants

Variables	*Minimum*	*Mean*	*Standard deviation*	*Maximum*	*Missing values*
Participants' Characteristics					
Age of participants	21	46	9.1	73	0
Number of years being a forest owner	2	13	7.9	40	0
Total number of forest parcels	1	2	1.0	4	0
Total forestland area (ha)	0.03	1.9	1.9	11.8	0
Cropland area (ha)	0	0.2	0.1	1	1
Total number of people in the participants' family	1	4	1.4	10	0
Number of females in the participants' family	0	2	1	6	0
Number of males in the participants' family	0	2	1	6	0
Number of people in the workforce in the participants' family	1	2	1	6	0
Number of people participating in forestry-related work in the participants' family	0	2	1	6	0
Level of annual income of participants	4	32	23.5	200	10
The contribution of crop income to annual income	0	50.1	24.1	100	0
The contribution of forestry income to annual income	0	24.9	19.9	90	0
The contribution of off-farm income to annual income	0	24.1	25.4	100	0
Forestland Management Objectives					
Owning forestland for generating forestry-related income	0	77.3	34.6	100	0
Owning forestland as an investment[5]	0	18.6	31.8	100	0
Owning forestland for creating natural landscape	0	3.8	13.3	100	0

With regard to family structure, there is an average of four people in each forest owner's family. Each family has at least one person (over 15 years of age) who is considered to be in the workforce. Interestingly, some families do not have any members participating in a forestry-related job. On average, two people in each family participate in forestry-related jobs.

With respect to annual income, ten participants were unwilling to mention their annual income. The average annual income of each participant is approximately 32 million

[5] Forestland is seen as an investment as the forest owners may sell the forestland if someone offers them a reasonable price.

VND/year[6]. Some participants can be considered high-income earners as their income exceeded 200 million VND/year.

On average, crop income contributes the most to the forest owners' annual income, accounting for approximately 50% of each forest owner's total annual income. Forestry and off-farm income contribute about 25% each to annual income. Interestingly, there were forest owners whose annual income was based completely on crop or off-farm income, but none claimed that their annual income was based completely on forestry-related income. A small proportion of forest owners also did not report any forestry-related income contribution in their income portfolios even though they owned forestland.

The results generally showed that the strongest motive in owning forestland was to generate income from forestry-related activities., followed by owning forest land as an investment and for creating natural landscapes.

The usable forms then were subdivided into afforestation and harvesting sets for modeling the respective afforestation and harvesting intensity. The criteria of this subdivision were explained in-depth in the afforestation and harvesting model section.

The modeling procedure involved two steps. The first step was to develop a series of one predictor models with the purpose of identifying candidate predictors for the final model, as described in Algorithm 1 in the previous chapter. The second step was to establish the final afforestation and harvesting models based on variables identified in step one. The second step involves a best subset selection procedure (Vincent Calcagno 2013) to determine the best final models.

5.5. Afforestation Intensity Models

The afforestation intensity models were developed according to the information on the forest owners' afforestation intensity since 2014. A group of 15 people was excluded from the sample because they already planted all of their forestland plots before 2014. An additional group of 22 people was also excluded because they only performed reforestation and did not

[6] 32 million VND is equivalent to 2048 NZD/year. At the time of writing the exchange rate is NZD1 = VND15,622.

have available forestland for afforestation. The final sample size for the afforestation intensity model is 281 participants from five different locations.

On average, the proportion of afforested area of these participants is 0.7. There are 71 and 186 participants who never afforested and afforested all of their forestland areas respectively. more than half of forested plots were planted with *Acacia mangium*. The rest was planted with other species including native species.

5.5.1. Variable Identification

One-predictor afforestation models are summarised in Tables 23, 24, 25, 26 and 27. The models are named according to the predictor's name. Each table is divided into three parts, which are described below:

1. The top rows of each table contain information about the maximum likelihood estimates of each coefficient, standard error and p-value. The p-value is computed via a Wald-statistics approximation. The standard error values are presented in parentheses
2. The Random Effects part of the table lists the global model variance and random effect variance.
3. The bottom row of each table lists the AIC values for the model.

Table 23 presents the one-predictor afforestation intensity models with respect to the personal characteristics of the forest owners. The result shows that the number of years of being a forest owner and the gender of the forest owner were not correlated with the afforestation intensity. These findings were consistent with those of the first study. The age and level of education of the forest owners were negatively correlated with the afforestation decision. The correlation with the age variable was consistent with the first study's result.

Table 23: A summary of afforestation intensity model with respect to personal characteristics

	age		ybfo		gen		edut	
Predictors	*Estimates*	*p*	*Estimates*	*p*	*Estimates*	*p*	*Estimates*	*p*
(Intercept)	0.56 (0.10)	**<0.001**	0.29 (0.04)	**<0.001**	0.35 (0.04)	**<0.001**	0.37 (0.04)	**<0.001**
Age	-0.01 (0.00)	**0.01**						
Number of years being forest owner (ybfo)			0.00 (0.00)	0.30				
Being female forest owners (gen)					-0.05 (0.04)	0.14		
Education level (edut)							-0.10 (0.04)	**<0.001**
Random Effects								
σ^2	0.16		0.16		0.16		0.16	
τ_{00}	0.02_{id}		0.03_{id}		0.02_{id}		0.02_{id}	
	0.00_{loc}		0.00_{loc}		0.00_{loc}		0.01_{loc}	
AIC	867.829		873.828		872.774		867.416	

Table 24 presents the one-predictor afforestation intensity model with respect to the family structure of the forest owners. The family size and family gender structure of the forest owners were not correlated with the afforestation intensity. These findings were consistent with those of the first study.

The number of people who were in the workforce in the forest owner’s family were negatively correlated with the afforestation decision. This impact was unexpected and suggested that the labour resources were allocated to non-forestry-related income generation activities. The number of people who were participating in forestry-related activities in the forest owner’s family was positively correlated with the afforestation intensity. The correlation of these predictors with the afforestation decision was different from the first study.

Table 24: A summary of afforestation intensity model with respect to family structure

	fco		fcom		fcof		fcow		fcopf	
Predictors	*Estimates*	*p*	*Estimates*	*p*	*Estimates*	*p*	*Estimates*	*p*	*Estimates*	*p*
(Intercept)	0.34 (0.06)	**<0.001**	0.32 (0.05)	**<0.001**	0.34 (0.05)	**<0.001**	0.44 (0.05)	**<0.001**	0.15 (0.05)	**<0.001**
Total number of people in forest owner's family (fco)	-0.00 (0.01)	0.80								
Total number of males in forest owner's family (fcom)			-0.00 (0.02)	0.99						
Total number of females in forest owner's family (fcof)					-0.01 (0.02)	0.76				
Total number of people in workforce in forest owner's family (fcow)							-0.04 (0.02)	**0.01**		
Total number of people participating forestry activities in forest owner's family (fcopf)									0.08 (0.02)	**<0.001**
Random Effects										
σ^2	0.16		0.16		0.16		0.16		0.16	
τ_{00}	0.03_{id}		0.03_{id}		0.03_{id}		0.02_{id}		0.02_{id}	
	0.00_{loc}		0.00_{loc}		0.00_{loc}		0.00_{loc}		0.00_{loc}	
AIC	874.826		874.887		874.795		869.142		850.963	

Table 25 presents one-predictor afforestation intensity models with regard to land assets and the forestland management objectives of the forest owners. The reason owning forestland for creating natural landscape was not corelated with afforestation intensity. This impact was inconsistent with the one in the first study. The total forestland area, total forestland plots and reason owning forestland for generating forestry income were positively correlated with afforestation intensity. These findings were consistent with those of the first study.

The total cropland areas and reason owning forestland as an investment were negatively correlated with afforestation intensity. These findings were inconsistent with those of the first study. In the first study these predictors were not correlated with the tree planting decision.

Table 25: A summary of afforestation intensity model with respect to the land assets and forestland management objectives

	tfl		tfp		crpha		ronlc		rogfi		roinv	
Predictors	*Estimates*	*p*	*Estimates*	*p*	*Estimates*	*p*	*Estimates*	*p*	*Estimates*	*p*	*Estimates*	*p*
(Intercept)	0.28 (0.04)	**<0.001**	0.23 (0.05)	**<0.001**	0.39 (0.04)	**<0.001**	0.33 (0.03)	**<0.001**	0.07 (0.04)	**0.04**	0.40 (0.03)	**<0.001**
Total forestland area (tfl)	0.02 (0.01)	0.05										
Total forestland plots (tfp)			0.04 (0.02)	**0.02**								
Total cropland area (crpha)					-0.33 (0.14)	**0.02**						
Reason owning forestland for creating natural landscape (ronlc)							-0.00 (0.00)	0.10				
Reason owning forestland for generating forestry income (rogfi)									0.00 (0.00)	**<0.001**		
Reason owning forestland as an investment (roinv)											-0.00 (0.00)	**<0.001**
Random Effects												
σ^2	0.16		0.16		0.16		0.16		0.16		0.16	
τ_{00}	0.02 $_{id}$		0.02 $_{id}$		0.02 $_{id}$		0.03 $_{id}$		0.01 $_{id}$		0.01 $_{id}$	
	0.00 $_{loc}$		0.00 $_{loc}$		0.00 $_{loc}$		0.00 $_{loc}$		0.00 $_{loc}$		0.00 $_{loc}$	
AIC	869.109		869.671		867.684		872.183		816.240		821.329	

Table 26 presents one-predictor afforestation intensity models with respect to the structure of the income of the forest owners. The natural logarithm of annual income and the contribution of off-farm income to annual income were negatively correlated with afforestation decisions. The contribution of forestry income was positively correlated with afforestation decision. The impacts of these predictors were consistent with those of the first study.

The contribution of crop income to annual income was not correlated with afforestation. This finding was inconsistent with the first study. The contribution of crop income to annual income was found to be significantly correlated with tree-planting decisions in the first study but were found to be not significant in the second study.

Table 26: A summary of afforestation intensity model with respect to the structure of income

	llin		crpin		frtin		othin	
Predictors	*Estimates*	*p*	*Estimates*	*p*	*Estimates*	*p*	*Estimates*	*p*
(Intercept)	0.56 (0.10)	**<0.001**	0.30 (0.05)	**<0.001**	0.24 (0.03)	**<0.001**	0.37 (0.03)	**<0.001**
Natural logarithm of annual income (llin)	-0.07 (0.03)	**0.01**						
Contribution of crop income to annual income (crpin)			0.00 (0.00)	0.56				
Contribution of forestry income to annual income (frtin)					0.00 (0.00)	**<0.001**		
Contribution of off-farm income to annual income (othin)							-0.00 (0.00)	**<0.001**
Random Effects								
σ^2	0.16		0.16		0.16		0.16	
τ_{00}	0.02 $_{id}$		0.03 $_{id}$		0.02 $_{id}$		0.02 $_{id}$	
	0.00 $_{loc}$		0.00 $_{loc}$		0.00 $_{loc}$		0.00 $_{loc}$	
AIC	846.588		874.553		862.208		865.367	

Table 27 summarises the effects of institutional support and market conditions on the afforestation intensity. The awareness of the forest owners about afforestation grants, technical support from forestry-extension workers, timber price and seedling cost[7] are positively correlated with the afforestation intensity. These findings are also in accord with the first study's results. The cost of buying fertilizer and the cost of planting trees were not significant in these models.

A limitation of the second survey was that the collected data was insufficient to fully model the tree-planting decisions using actual market price of timber and the cost of seedlings as initially planned. In the questionnaire, the interviewees were required to provide the market price of timber and seedlings. However, of the 281 forest owners who participated in the survey, only 75 people provided the actual market price of timber. Similarly, only a fraction of the forest owners provided the cost of seedlings and fertilizer. Despite this limitation, the findings suggest that the timber price and the cost of buying seedlings factors in the decisions of the forest owners.

[7] These are categorical predictors. The detail of these predictor can be found in Appendix 1.

Table 27: A summary of the afforestation intensity mixed-effects regressions with respect to the institutional support and market condition

	grntaw		**sup**		**tpr**		**seed**		**fert**		**cplnt**	
Predictors	*Estimates*	*p*	*Estimates*	*p*	*Estimates*	*p*	*Estimates*	*p*	*Estimates*	*p*	*Estimates*	*p*
(Intercept)	0.24 (0.03)	**<0.001**	0.20 (0.02)	**<0.001**	0.30 (0.03)	**<0.001**	0.27 (0.04)	**<0.001**	0.31 (0.03)	**<0.001**	0.32 (0.04)	**<0.001**
The awareness of the forest owners about government grant (grntaw)	0.13 (0.04)	**<0.001**										
Receiving technical support from forestry extension workers (sup)			0.23 (0.03)	**<0.001**								
Including timber price in the decision (tpr)					0.16 (0.05)	**<0.001**						
Including cost of seedlings in the decision (seed)							0.17 (0.04)	**<0.001**				
Including cost of fertilser in the decision (fert)									0.07 (0.04)	0.11		
Including cost of planting in the decision (cplnt)											0.01 (0.04)	0.76
Random Effects												
σ^2	0.16		0.16		0.16		0.16		0.16		0.16	
τ_{00}	0.02_{id}		0.01_{id}		0.02_{id}		0.02_{id}		0.03_{id}		0.03_{id}	
	0.00_{loc}		0.00_{loc}		0.00_{loc}		0.00_{loc}		0.00_{loc}		0.00_{loc}	
AIC	859.657		817.528		863.372		855.873		862.547		863.209	

In summary, the factors affecting the afforestation intensity of the forest owners, according to the one-predictor models, are as follows:

1. Age of the forest owners.
2. Level of education of the forest owners.
3. Total number of people in workforce in the forest owners' family.
4. Total number of people participating in forestry activities in the forest owners' family.
5. Total forestland area.
6. Total forestland plots.
7. Total cropland area of the forest owners.
8. Reason owning forestland for generating forestry income.
9. Reason owning forestland as an investment.
10. Natural logarithm of annual income.
11. The contribution of forestry income to annual income.
12. The contribution of off-farm income to annual income.
13. Awareness about government tree-planting subsidy program.
14. Receiving technical support from forestry extension workers.
15. Timber price.
16. Cost of buying seedlings.

5.5.2. Variance Inflation Factor (VIF)

After identifying the ideal set of predictors, it is necessary to detect collinearity between the predictors before moving on to step two. According to Harrison et al. (2018) collinearity can be detected by determining the correlation matrices between the raw explanatory variables or calculating the *Variance Inflation Factor (VIF)* of each predictor. The latter method is particularly useful in this study because it allows random effects to be included. All VIF analyses were carried out using the *car* R package (Fox and Weisberg 2011)

The third column in Table 28 presents the VIF values of the identified predictors. One may observe that the VIF values of reasons owning forestland for generating forestry income and reason for owning forestland as an investment predictor are greater than 10. The other major predictors have VIF values ranging from 1.1 to 2.

This study proposes three different alternatives for excluding collinear predictors. The first option would be excluding the reason owning forestland for generating forestry income predictor. The second option is identical to the first but excludes the reason for owning forestland as an investment predictor. The last option is the most practical choice with respect to a management point of view. The nominated predictors for this option are selected for its reliability and validity in the forestry management system in Vietnam. The advantage of this approach is that the final model can be easily adapted for use by the government. The detail of these predictors' VIF value is presented in Table 28.

Table 28: VIF value of afforestation predictors

No.	Predictors	All predictors	Option 1	Option 2	Option 3
1	Age of the forest owner	1.2	1.2	1.2	1.2
2	Education level	1.2	1.2	1.2	1.2
3	Total people in workforce in the forest owners' family	1.7	1.7	1.8	1.2
4	Total people participating in forestry activities in the forest owners' family	2.0	2.0	2.0	
5	Total forestland area	1.7	1.7	1.7	1.6
6	Total forestland plot	1.8	1.8	1.8	1.6
7	Total cropland area	1.2	1.2	1.2	1.2
8	Natural logarithm of annual income	1.5	1.5	1.5	1.3
9	Contribution of forestry income to annual income	1.8	1.8	1.7	
10	Contribution of off-farm income to annual income	1.8	1.7	1.7	
11	Reason owning forestland for generating forestry income	10.7		2.0	
12	Reason owning forestland as an investment	10.7	2.0		
13	Timber price	1.1	1.1	1.1	
14	Cost of buying seedling	1.2	1.2	1.2	
15	Technical support from government extension workers	1.6	1.6	1.6	
16	Awareness about government subsidy grant for establishing forests	1.6	1.6	1.6	1.2

5.5.3. Afforestation Intensity Model

The final model was chosen based on the best subset selection procedure that was described in the previous chapter. The procedure generates all possible combinations of predictors from each input set, which were then used as the input to create models. These models were sorted according to their AIC value and the model with lowest AIC value was considered as the best model.

Table 29 summarises the key statistics of the best model from each of the three sets of nominated variables. The best models are named according to their input set. What stands out in the table is that the variation between the locations in all models is essentially zero. Regarding the AIC values, there is a notable difference between the first two models and the last one. The difference of AIC value between the first and second model in comparison with the last model are 59 and 61 units of AIC, respectively.

Comparing the first two models, it can be seen that they use an equal number of predictors. The single most striking difference is that there is no significant difference with respect to AIC value. The difference between the two models is the use of the reason for owning forestland for generating forestry income as a predictor in option 1 and the reason for owning forestland as an investment as a predictor in the option 2. Hence, it can be said that those predictors are virtually interchangeable despite the fact that they have an opposite sign in the models.

Table 29 also reveals the overall pattern of the data interaction in the three best models. It can be observed that the variables utilised in these models directly or indirectly belong into four different categories. These categories are forestland management objectives, labour distribution, institutional factors and market condition.

With respect to forestland management objectives, the first two models explicitly use either the reason owning forestland as an investment or the reason owning forestland for generating forestry income as predictor. These predictors negatively and positively affect the afforestation decisions in the first and second model respectively. Obviously, the forest owners who own the forestland for generating forestry income have positive behaviour towards afforestation because it was the reason that they own the forestland.

The investment forest owners have less incentives to plant trees because to establish a forest requires additional expenses on planting trees and maintaining the forests. Furthermore, there are risks that can affect their expected returns such as bushfire, decrease in timber price etc.

These challenges could lead to negative behaviour of the investment forest owners towards tree planting.

It also can be observed that the AIC value of the model significantly increases when these two variables are not in the models. Therefore, it can be concluded that these variables have notable influence on the afforestation decision.

The last model does not directly include the reason for owning forestland predictors. The total forestland plots of the forest owners may be considered as an indicator of forestland management objectives. It has a significant positive impact on the afforestation decision. A possible explanation is that the investment forest owners may tend to own less forestland plots in comparison with owners who want to generate forestry income. Buying a piece of forestland requires a substantial amount of money. Thereby, the opportunity cost of buying forestland might be a reason that inhibits their incentives of owning additional forestland plots. Meanwhile, owning more forestland plots can provide a better opportunity to increase income for the owners who want to generate forestry income.

With regard to the institutional factors and market conditions, the first two models use the cost of buying seedlings and technical support from government extension workers as indicators. Due to the exclusion of these predictors from the set of nominated predictors, the last model uses the tree planting grant awareness variable. Interestingly, supplying seedlings and technical support are, in fact, key components of the grant package that the grant recipients will receive if they participate in the grant scheme. Thus, the tree planting grant awareness predictor, to some extent, indirectly implies that the forest owners take seedling cost and technical support into consideration.

Predictors related to labour distribution include the age of the forest owners, the level of education of the forest owners, the total number of family members in the workforce, the total number of family members participating in forestry related activities and the total cropland area. Of these variables, the age, the total number of family members in the workforce and the total number of family members participating in forestry related activities explicitly represent the labour distribution in the forest owner's family.

Meanwhile, the total cropland area and level of education of the forest owners indirectly describe the labour allocation practices of the forest owners. To put it in another way, if labour is allocated to agricultural practice, it results in fewer available labour resources for doing forestry activities. For the level of education predictor, a higher level of education might result in a better chance of finding off-farm jobs which can provide better income returns.

In the first two models, the sign of the total cropland area and the total number of family members participating in forestry related activities predictor in the models are negative and positive respectively. A possible explanation for this is that tree planting is a labour-intensive activity. A larger number of people who can participate in forestry-related activities is an indicator of greater incentives to plant trees. In contrast, the negative sign of the total cropland area predictor implies that if the forest owners have more crop land area, they tend to allocate their labour resources to agricultural practice instead of forestry. This may be due to the fact that revenue derived from forestry-related activities takes a longer time to show significant returns. This factor, in combination with the low level of income of the forest owners, could explain why the forest owners focus on agriculture when their cropland area is relatively large.

Logically, it is expected that a greater number of people in the forest owner's families in the workforce would be correlated with greater incentives for planting trees. However, contrary to expectations, the number of people in the forest owner's families in the workforce is negatively correlated with the afforestation intensity. A possible explanation for this is that these people may be allocated to other jobs including agriculture and off-farm activities. Hence, an increase in the number of working people in the forest owner's family does not correspond to an increase in the afforested area.

In the last model, the age and level of education can be considered as predictors representing the labour distribution. The age and level of education predictors are negatively correlated with the afforestation decision. It is understandable because tree planting and forest maintenance are physically intensive activities. Hence, the aged and educated forest owners may allocate their daily activities to more profitable jobs.

Figure 16 graphically shows the estimated relative importance of model terms, which are normalized and sum up to one. The red vertical line drawn at 0.8 is an arbitrary division that is used as a cutoff differentiating the important and the not so important terms. It can be seen that the orders of term importance with respect to option one and two are similar. The forestland management objective including reason owning forestland as an investment (Roinv) and reason owning forestland for generating forestry income (Rogfi) are the second most important predictor in the set. It confirms the aforementioned discussion about the role of these predictors in the models.

It can be observed that technical support and cost of seedlings are the top five predictors in the first two sets of nominated variables, and that, conversely, the awareness about the grant plays the least important role. However, when the cost of seedlings and technical support predictors are excluded from the set, the forest owner's awareness about the grant takes the top position.

Taking into consideration the practicalities of applying the final model in the provincial forestry management system, the third model is recommended. The model can be rewritten in an equation of the following form:

$$\textit{Afforestation intensity} = 0.47 + 0.04\,tfp - 0.005\,age - 0.09\,edut - 0.34\,crpha + 0.15\,grntaw + e$$

Where:

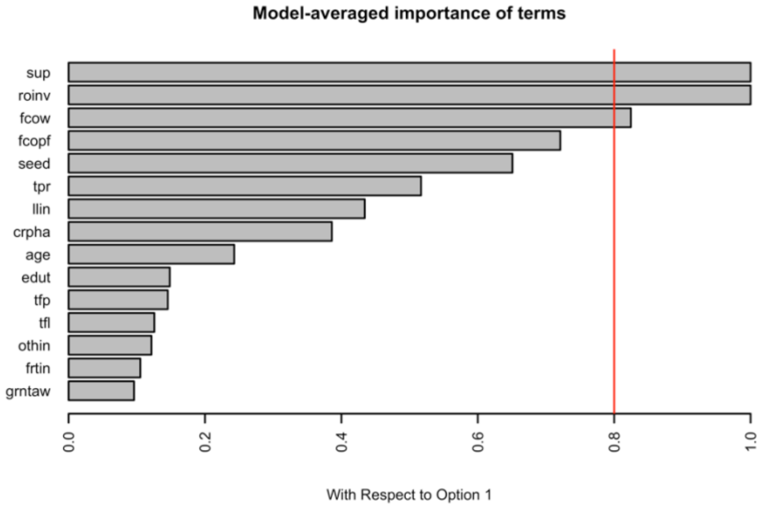
- crpha: total cropland area of the forest owners.
- age: the age of the forest owners.
- edut: education level of the forest owners.
- grntaw: Awareness about government susdidy grant for establishing forests.
- tfp: total forestland plot of the forest owners.
- *e:* the residuals that represents other factors affecting the decision.

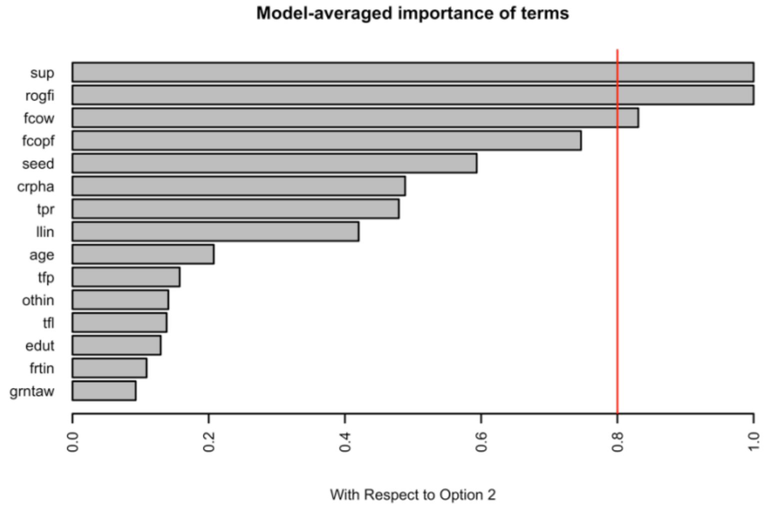
The model can be interpreted as ceteris paribus effect of the total cropland area of the forest owners, the age of the forest owners, the level of education of the forest owners, the awareness of the forest owners about government tree planting subsidy grant, the total forest plots of the forest owners and the residuals which are contained in *e*.

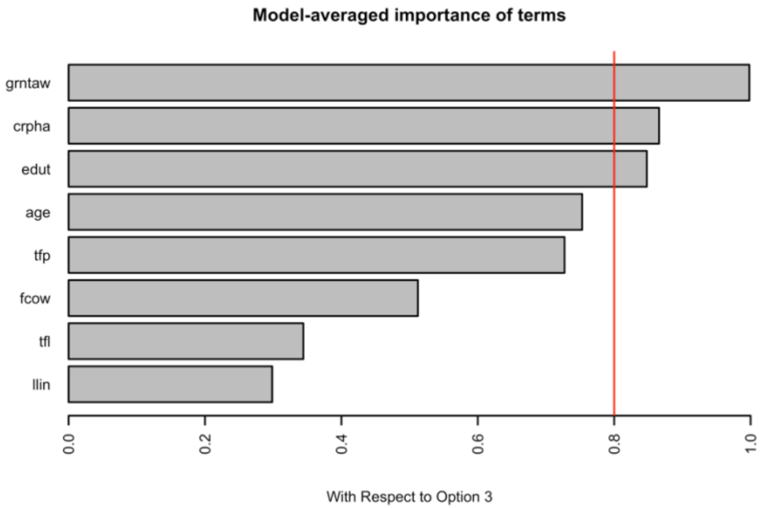
A unit increase in total cropland area and age of the forest owners leads to a decrease in the average afforestation intensity of the forest owners by approximately 0.34 and 0.005 respectively. The afforestation intensity of the forest owners who have completed high school or higher education tend to be lower than owners possessing an intermediate school degree or a lower level of education by approximately 0.09. The afforestation intensity of the forest owners who knew about the tree planting subsidy grant tend to be higher than the ones who did not know about the grant by approximately 0.15. An increase of one forestland plot leads to the increase in average afforestation intensity of the forest owners by approximately 0.04.

Table 29: Summary of the best afforestation intensity models

	Option.1			*Option.2*			*Option.3*		
Predictors	*Estimates*	*std.Error*	*P*	*Estimates*	*std.Error*	*P*	*Estimates*	*std.Error*	*P*
(Intercept)	0.1421	0.0974	0.14	-0.0356	0.1052	0.74	0.4703	0.1031	<0.001
Forestland management objectives									
Reason owning forestland as an investment	-0.0019	0.0005	<0.001						
Reasoning owning forestland for generating forestry income				0.0019	0.0005	<0.001			
Total forestland plots							0.0369	0.0173	0.03
Labour distribution									
Age of the forest owners							-0.0045	0.002	0.02
Education level							-0.0888	0.0346	0.01
Natural logarithm of annual income	0.0577	0.0266	0.03	0.0577	0.0265	0.03			
Total number of people in workforce in the forest owners' family	-0.0515	0.0171	<0.001	-0.0508	0.0171	<0.001			
Total number of people participating in forestry activities in the forest owners' family	0.0472	0.0186	0.01	0.0465	0.0183	0.01			
Total cropland area	-0.2081	0.1213	0.09	-0.2349	0.1207	0.05	-0.3398	0.1373	0.01
Institutional factors									
Receiving technical support from forestry extension workers	0.1785	0.0319	<0.001	0.1791	0.0317	<0.001			
Awareness about government subsidy grant for establishing forests							0.1464	0.0345	<0.001
Market conditions									
Cost of buying seedlings	0.0842	0.0364	0.02	0.0795	0.0363	0.03			
Random Effects									
Global model variance		0.16			0.16			0.16	
Person level variance		0.00			0.00			0.01	
Location level variance		0.00			0.00			0.00	
Observations		730			730			730	
AIC		747.830			745.573			806.381	







Legend:	
age: Age of the forest owner	**frtin**: Contribution of forestry income to annual income
edut: Education level	**othin**: Contribution of off-farm income to annual income
fcow: Total people in workforce in the forest owners' family	**rogfi**: Reason owning forestland for generating forestry income
fcopf: Total people participating in forestry activities in the forest owners' family	**roinv**: Reason owning forestland as an investment
tfl: Total forestland area	**tpr**: Timber price
tfp: Total forestland plot	**seed**: Cost of buying seedling
crpha: Total cropland area	**sup**: Technical support from government extension workers
llin: Natural logarithm of annual income	**grntaw**: Awareness about government subsidy grant for establishing forests

Figure 16: Afforestation model – Averaged Importance of Terms

5.5.4. Summary

This section presents a summary of the process used to create the afforestation model. Firstly, the analysis has identified the three best candidate models from the three different sets of predictors using AIC values as a benchmark. The predictors included in these models directly or indirectly belong to four different categories:

1. Forestland management objectives: reason owning forestland for generating income from forestry related activities, reason owning forestland as an investment, total forest plots of the forest owners.
2. Institutional factors: technical support from government extension workers and awareness about government subsidy grant for establishing forests.
3. Market conditions: cost of buying seedlings.
4. Labour distribution in the forest owners' family: age of the forest owners, the level of education of the forest owners, total number of family members in the workforce, total number of family members participating in forestry related activities and total cropland area of the forest owners.

Secondly, the reason owning forestland as an investment and reason owning forestland for generating forestry income show high multicollinear correlation with each other. If one of these variables is excluded from the set of predictors there is no presence of high multicollinearity among predictors. Additionally, it was observed that these variables play a relatively equal role in the models with respect to the AIC value.

In conclusion, the following model is recommended to be applied to the provincial forestry management system:

$$\textit{Afforestation intensity} = 0.47 - 0.34\,crpha - 0.005\,age - 0.09\,edut + 0.15\,grntaw + 0.04\,tfp + e$$

The model can be interpreted as the relationship between the afforestation intensity and the combined ceteris paribus effect of the total cropland area of the forest owners, the age of the forest owners, the level of education of the forest owners, the awareness of the forest owners about government tree planting subsidy grant, the total forest plots of the forest owners and the residuals which are contained in *e*.

5.6. Harvesting Model

The desired output of the harvesting intensity model is the proportion of harvested area versus the total available forests for harvesting. Therefore, the forest owners who never planted trees were excluded from the sample. With respect to the planted forest plots, the dominant planted species was *Acacia mangium*. A small proportion of forestland area was covered by mixed species. Therefore, forest plots that were planted with mixed species were also excluded from the sample. The final sample size of the harvest intensity model consisted of 105 participants from five locations.

According to Sein and Mitlöhner (2011), the rotation of *Acacia mangium* in Vietnam varies from five to seven years depending on the aims and means of the forest owners. It suggests that the optimum tree age for harvesting peaks from the tree age of five to seven. Figure 17 presents the distribution of optimum tree age for harvesting provided by the research participants. It can be observed that about half of the forest owners in harvest sample pool decided to harvest at a tree age of seven. Additionally, the figure indicates that the number of people planning to harvest their forests increases and decreases before and after tree age of seven.

Because of this, it is more interesting if model can describe the harvest intensity before and after optimum tree age for harvesting. Taking this into consideration, this research used the deviation of tree age as a predictor in the harvesting intensity models instead of using the actual tree age. The deviation of tree age in this research is defined as the modulus between actual age of standing or harvested trees and average of optimum tree age for harvesting that is provided by the research participants.

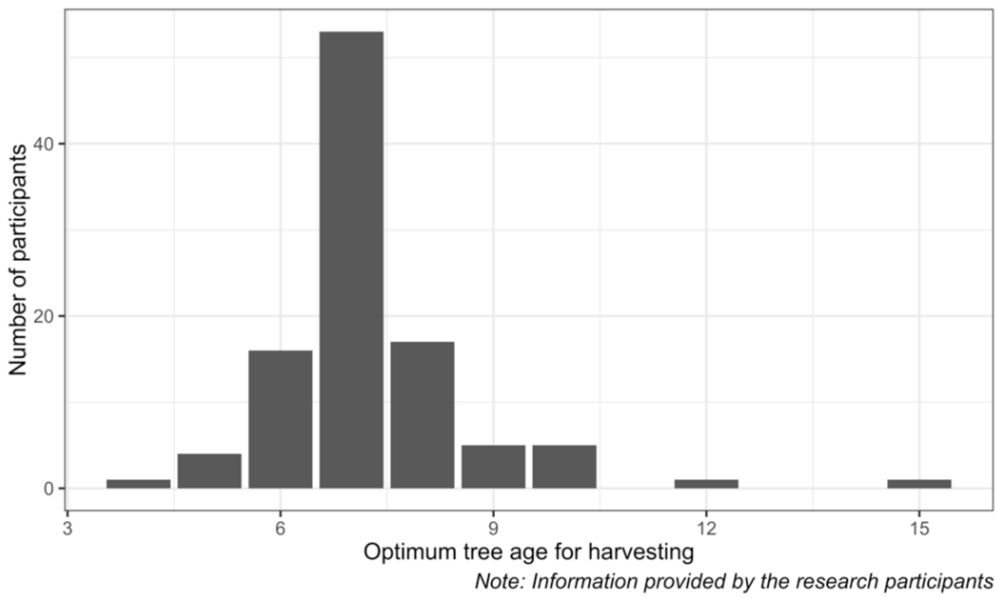


Figure 17: Distribution of optimum harvest tree age

In summary, the response of the model is the proportion of harvested area versus the total area of available forests for harvesting, with specific regard to *Acacia mangium*. On average, the proportion of harvested area for these participants is 0.25. There were 42 participants who did not harvest their forests and 7 who harvested all of their forests respectively.

On average, the forest owners sold their forests for price of 1 million VND/cubic metre[8] . The lowest and highest price were 0.3 and 1.7 million VND respectively. There were 34 participants who could not provide the timber price.

5.6.1. Variable Identification

Table 30 presents the one-predictor harvest intensity models with respect to the personal characteristics of the forest owners. The result shows that the age of the forest owners and level of education of the forest owners were not correlated with the harvest intensity. The number of years being a forest owner was positive correlated with the harvest intensity. In contrast, being a female forest owner was negatively correlated with harvest intensity.

[8] 1 Million VND is equivalent to 64 NZD. At the time of writing the exchange rate is NZD1 = VND15,622. It is a negotiated price between tree-farmers and traders with respect to standing trees.

Table 30: A summary of harvest intensity models with respect to personal characteristics

	age		ybfo		gen		edut	
Predictors	*Estimates*	*p*	*Estimates*	*p*	*Estimates*	*p*	*Estimates*	*p*
(Intercept)	0.030 (0.106)	0.78	0.083 (0.058)	0.16	0.220 (0.045)	**<0.001**	0.189 (0.046)	**<0.001**
Age	0.003 (0.002)	0.12						
Number of years being forest owner (ybfo)			0.006 (0.002)	**0.01**				
Gender (gen)					-0.100 (0.039)	**0.01**		
Education level (edut)							-0.011 (0.042)	0.79
Random Effects								
σ^2	0.13		0.12		0.12		0.13	
τ_{00}	0.00_{id}		0.00_{id}		0.00_{id}		0.00_{id}	
	0.01_{loc}		0.01_{loc}		0.01_{loc}		0.01_{loc}	
AIC	278.682		274.551		274.655		281.069	

Table 31 presents the one-predictor harvest intensity models with respect to the family structure of the forest owners. The total number of people in the forest owner's family is positively correlated with the harvest intensity. A possible explanation for this might be that the forest owners who have a greater family size might also be faced with greater financial pressure for sustaining their family in comparison with the ones that have a smaller family size. This might be a strong motivator for the harvesting decisions.

The remaining predictor, including the total number of males and females, the total number of people in the workforce and the total number of people participating in forestry activities in the forest owners' family were not correlated with the harvesting intensity.

Table 31: A summary of harvest intensity models with respect to family structure

	fco		fcom		fcof		fcow		fcopf	
Predictors	*Estimates*	*p*	*Estimates*	*p*	*Estimates*	*p*	*Estimates*	*p*	*Estimates*	*p*
(Intercept)	0.062 (0.078)	0.43	0.121 (0.065)	0.06	0.120 (0.060)	**0.05**	0.198 (0.067)	**<0.001**	0.163 (0.064)	**0.01**
Total number of people in forest owner's family (fco)	0.026 (0.014)	0.05								
Total number of males in forest owner's family (fcom)			0.026 (0.020)	0.18						
Total number of females in forest owner's family (fcof)					0.029 (0.019)	0.13				
Total number of people in workforce in forest owner's family (fcow)							-0.005 (0.018)	0.79		
Total number of people participating forestry activities in forest owner's family (fcopf)									0.009 (0.020)	0.66
Random Effects										
σ^2	0.12		0.13		0.13		0.13		0.13	
τ_{00}	0.00_{id}		0.00_{id}		0.00_{id}		0.00_{id}		0.00_{id}	
	0.01_{loc}		0.01_{loc}		0.01_{loc}		0.01_{loc}		0.01_{loc}	
AIC	277.560		279.418		278.897		281.070		280.950	

Table 32 presents the one-predictor harvest intensity models with respect to land assets and forestland management objectives. The reason for owning forestland for generating forestry income was positively correlated with harvest intensity. Obviously, the forest owners who want to generate income from forestry activities would harvest their forests to achieve their objectives.

In contrast, the reason owning forestland as an investment was negatively correlated with harvest intensity. According to Dennis (1989), harvesting behaviour is influenced by the level of exogenous income. An increase in the level of exogenous income leads to a reduction in the marginal utility of income derived from harvesting forest. Therefore, it can be concluded that the forest owners’ investment income is generally secured from other sources that were not from forestry activities.

The remaining predictors including the total forestland area of the forest owners, the total number of forestland plots of the forest owners, the total cropland area of the forest owners, and the reason for owning forestland for creating natural landscape, were not correlated with the harvest intensity.

Table 32: A summary of harvest intensity models with respect to the land assets and forest management objectives

	tfl		**tfp**		**crpha**		**ronlc**		**rogfi**		**roinv**	
Predictors	*Estimates*	*p*	*Estimates*	*p*	*Estimates*	*p*	*Estimates*	*p*	*Estimates*	*p*	*Estimates*	*p*
(Intercept)	0.189 (0.052)	**<0.001**	0.219 (0.069)	**<0.001**	0.223 (0.050)	**<0.001**	0.183 (0.042)	**<0.001**	0.044 (0.084)	0.60	0.202 (0.042)	**<0.001**
Total forestland area (tfl)	-0.001 (0.009)	0.88										
Total forestland plots (tfp)			-0.013 (0.021)	0.52								
Total cropland area (crpha)					-0.174 (0.148)	0.24						
Reason owning forestland for creating natural landscape (ronlc)							0.000 (0.002)	0.82				
Reason owning forestland for generating forestry income (rogfi)									0.002 (0.001)	0.06		
Reason owning forestland as an investment (roinv)											-0.002 (0.001)	**0.02**
Random Effects												
σ^2	0.13		0.13		0.13		0.13		0.12		0.12	
τ_{00}	0.00_{id}		0.00_{id}		0.00_{id}		0.00_{id}		0.00_{id}		0.00_{id}	
	0.01_{loc}		0.01_{loc}		0.00_{loc}		0.01_{loc}		0.01_{loc}		0.01_{loc}	
AIC	281.112		280.728		279.826		281.085		277.478		275.884	

Table 33 presents the one-predictor harvest intensity models with respect to the structure of annual income. The contribution of crop income to annual income was negatively correlated with the harvest intensity. The impact of this predictor confirms the aforementioned prescription of Dennis (1989) about the level of exogenous income.

The remaining predictors including the natural logarithm of annual income, the contribution of forestry income to annual income, the contribution of off-farm income to annual income were not correlated with harvest intensity.

Table 33: A summary of harvest intensity models with respect to the structure of income

	llin		crpin		frtin		othin	
Predictors	*Estimates*	*p*	*Estimates*	*p*	*Estimates*	*p*	*Estimates*	*p*
(Intercept)	0.157 (0.108)	0.15	0.280 (0.052)	**<0.001**	0.165 (0.055)	**<0.001**	0.170 (0.043)	**<0.001**
Natural logarithm of annual income (llin)	0.007 (0.030)	0.82						
Contribution of crop income to annual income (crpin)			-0.002 (0.001)	**0.02**				
Contribution of forestry income to annual income (frtin)					0.001 (0.001)	0.61		
Contribution of off-farm income to annual income (othin)							0.001 (0.001)	0.39
Random Effects								
σ^2	0.12		0.13		0.13		0.13	
τ_{00}	0.00_{id}		0.00_{id}		0.00_{id}		0.00_{id}	
	0.01_{loc}		0.00_{loc}		0.01_{loc}		0.01_{loc}	
AIC	266.028		276.783		280.877		280.432	

Table 34 presents the effects of tree biological, personal and institutional factors and market conditions on the harvest intensity. Government regulation of forest harvesting was not correlated with harvest intensity. The harvest intensity is positively correlated with the increase in timber price. The deviation of tree age was negatively correlated with harvest intensity, which was to be expected as the harvest magnitude reaches its peak when the tree age is seven.

The forest owners who consider forests as type of financial saving were less likely to harvest the forests. The forest owners also included the cost of harvesting in their decision. Unfortunately, the provided information about the actual cost of harvesting was not sufficient to fully model harvest intensity.

Table 34: A summary of harvest intensity models with respect to tree biological, personal and institutional factors and market conditions

	devage		fsave		govhar		char		ltpr	
Predictors	*Estimates*	*p*	*Estimates*	*p*	*Estimates*	*p*	*Estimates*	*p*	*Estimates*	*p*
(Intercept)	0.257 (0.037)	**<0.001**	0.231 (0.030)	**<0.001**	0.191 (0.041)	**<0.001**	0.144 (0.055)	**0.01**	-1.168 (0.429)	**0.01**
Deviation of tree age (devage)	-0.056 (0.016)	**<0.001**								
Forests as type of financial saving (fsave)			-0.093 (0.040)	**0.02**						
Government regulation on harvesting forest (govhar)					-0.032 (0.072)	0.65				
Including cost of harvesting in the decision (char)							0.075 (0.040)	0.06		
Natural logarithm of timber price (ltpr)									0.208 (0.063)	**<0.001**
Random Effects										
σ^2	0.12		0.12		0.13		0.12		0.15	
τ_{00}	0.00_{id}		0.00_{id}		0.00_{id}		0.00_{id}		0.00_{id}	
	0.00_{loc}		$0.00_{id.1}$		0.01_{loc}		0.01_{loc}		0.00_{loc}	
AIC	269.963		272.342		280.552		271.372		221.904	

5.6.2. Variance Inflation Factor (VIF)

The third column in Table 35 presents the VIF value of the identified predictors. One may observe that none of these predictors had VIF greater than 10. However, the VIF value of the reason owning forestland for generating forestry income and the reason owning forestland as an investment predictors were substantially higher than the others. Taking this point into consideration, this study proposes three different alternatives for excluding collinear predictors.

The first option consists of excluding the reason owning forestland for generating forestry income predictor. The second option excludes the reason owning forestland as an investment predictor. Option three is the most practical choice with respect to a management point of view. The nominated predictors for this optin were selected for their reliability and validity in the forestry management system in Vietnam. The advantage of this approach is that the final model can be easily adapted for use by the government. The details of these predictors' VIF are presented in Table 35.

Table 35: VIF value of harvest intensity predictors

No.	Predictors	All predictors	Option 1	Option 2	Option 3
1	Number of years being forest owners	1.4	1.4	1.4	1.3
2	Gender of the forest owners	1.2	1.3	1.2	1.1
3	Family size of the forest owners	1.3	1.3	1.3	1.2
4	Contribution of crop income to annual income	1.2	1.2	1.2	
5	Reason owning forestland for generating forestry income	6.4		1.2	
6	Reason owning forestland as an investment	6.2	1.2		
7	Deviation of tree age	1.4	1.4	1.1	1.
8	Considering forests as financial saving	1.3	1.3	1.2	
9	Including cost of harvesting in the decision	1.4	1.4	1.4	
10	Natural logarithm of timber price	1.2	1.2	1.2	1.0

5.6.3. Harvest Intensity Model

Three sets of candidate predictors were used to develop the final harvest intensity model using the best subset selection method. The worst model AIC values of option 1, 2 and 3 were 223, 225 and 220 respectively. The best models generated from option 1 and 2 had very similar AIC values.

Table 36 presents the key statistics of the best model from each of the three sets of nominated predictors. The best models are named according to their input set. What stands out in the table is that the AIC value between the best models from different sets is similar. It can be said that these models play an equal role in explaining the harvest intensity of the forest owners. All models used the deviation of tree age and the natural logarithm of the timber price predictor.

The best model derived from option 1 and 2 included the cost of harvesting predictor in the model. The presence of the cost of harvesting in the model suggests that the forest owners were not only thinking of total revenue from the harvesting of forests but also the profit margin of their forestry activities. This finding is consistent with the first study.

In contrast, the best model derived from option three used the gender of the forest owners in the model. This result indicates that being a female forest owner is correlated with a lower harvest intensity. A possible explanation for this is that women tend to spend more time on domestic activities. Consequently, they might be less inclined to actively seek out market information with respect to timber products.

Figure 18 graphically presents the estimated relative importance of model terms, which are normalized and sum up to one. It can be seen that the three most important terms are the natural logarithm of timber price, deviation of tree age and gender of the forest owners.

Table 36: Summary of the best harvest intensity model

	Option 1 and 2			**Option 3**		
Predictors	*Estimates*	*std.Error*	*P*	*Estimates*	*std.Error*	*P*
(Intercept)	-1.135	0.463	0.01	-0.806	0.427	0.06
Cost of harvesting	0.104	0.057	0.07			
Deviation of tree age	-0.069	0.028	0.01	-0.070	0.028	0.01
Natural logarithm of timber price	0.205	0.066	<0.001	0.169	0.061	0.01
Female forest owners				-0.104	0.058	0.07
Random effects						
Global model variance		0.14			0.15	
Personal level variance		0.00			0.00	
Location level variance		0.00			0.00	
AIC		208.73			208.811	

Taking into consideration the practicalities of applying the final model in the provincial forestry management system, model option three is recommended. The model can be written in equation form as:

$$Harvest\ intensity = -0.8 - 0.07*devage + 0.17*lacttpr - 0.1*gen + e,$$

Where:

- devage - absolute deviation of tree age.
- lacttpr - natural logarithm of timber price.
- gen - gender of the forest owners (takes values of 0 and 1 with respect to female and male forest owners, respectively).
- *e:* the residuals that represent other factors affecting the decision.

The model can be interpreted as ceteris paribus effect of the deviation of tree age, natural logarithm of timber price and the gender of the forest owners on harvest intensity. A unit of difference between the actual tree age and a tree age of seven results in reduction of harvest intensity by approximately 0.07. An increase in a unit of natural logarithm of timber price results in an increase of harvest intensity by approximately 0.17. Lastly, being female forest owner results in reduction of harvest intensity by approximately 0.1.

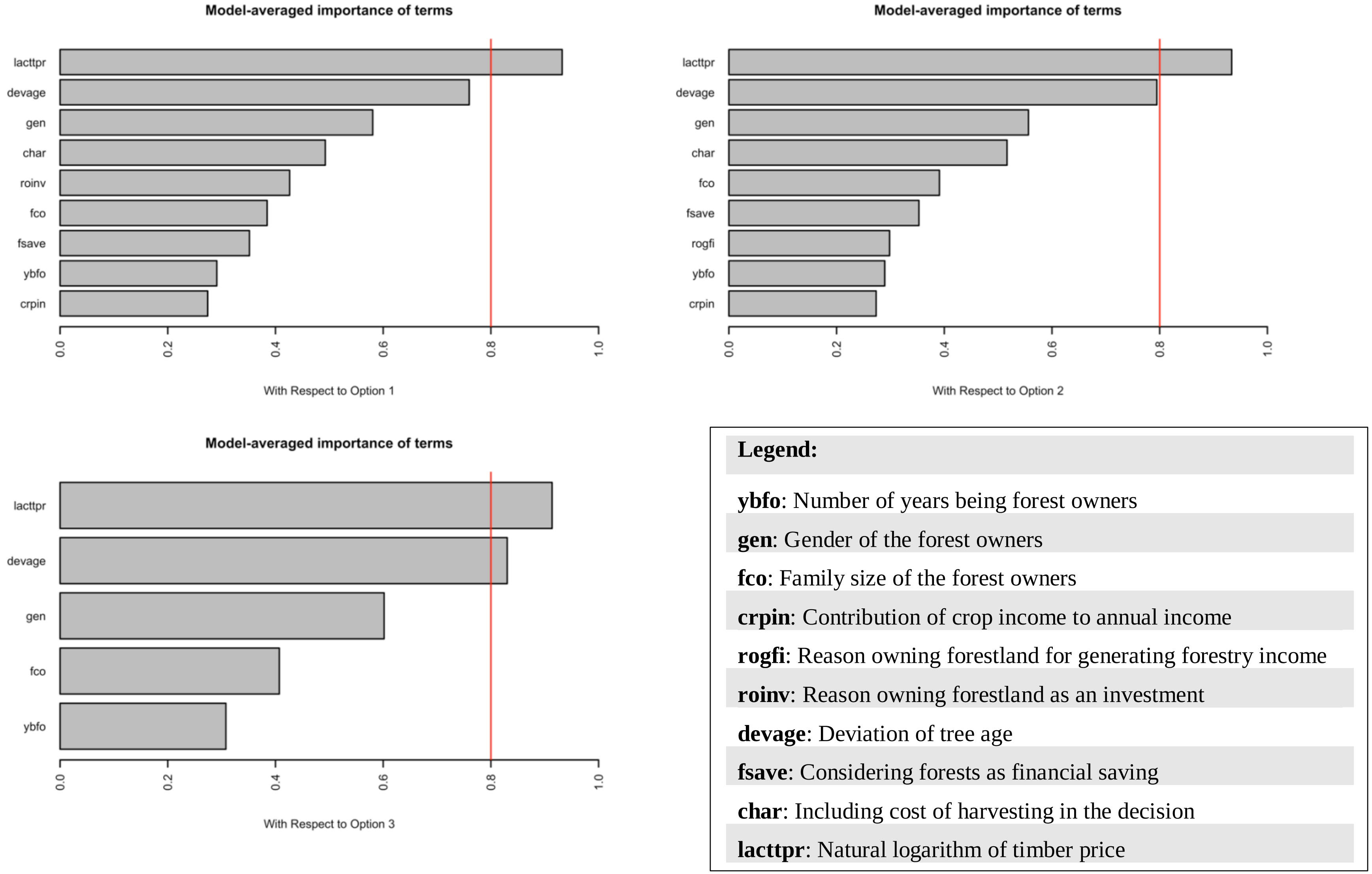


Figure 18: Harvest model-averaged importance of terms

5.6.4. Summary

The purpose of this analysis was to model the harvest intensity of the forest owners with respect to *Acacia mangium* in Thai Nguyen province, Vietnam. In summary, we identified ten predictors affecting the harvesting decisions of the forest owners:

1. Number of years being a forest owner.
2. Gender of the forest owners.
3. Total number of people in the forest owners' family.
4. Considering forests as a type of financial saving.
5. Contribution of crop income to annual income.
6. Reason for owning forestland for generating forestry income.
7. Reason for owning forestland as an investment.
8. Deviation of tree age to the mean of the expected tree age for harvesting.
9. Cost of harvesting.
10. Natural logarithm of timber price.

Of these predictors, the reason owning forestland for generating forestry income and the reason owning forestland as an investment showed high multicollinearity with each other. If one of these predictors is excluded from the set of predictors, there is no presence of high multicollinearity among predictors.

Three sets of predictors were used to develop optimal multiple-predictor models using the best subset selection procedure. AIC was used as a benchmark for identifying the best model. Two best models were identified, as there was no significant difference with respect to the AIC value between the two models. Taking into consideration the practicalities of applying the final model in the provincial forestry management system, the third model was recommended. The model can be interpreted as ceteris paribus effect of the deviation of tree age, natural logarithm of timber price and the gender of the forest owners on the harvest intensity of *Acacia mangium*. The model can be written in equation form as:

$$Harvest\ intensity = -0.8 - 0.07*devage + 0.17*lacttpr - 0.1*gen + e$$

5.7. Conclusion

The second study identified factors affecting afforestation and harvest intensity of the forest owners. With respect to the harvest intensity model, focus was solely on modelling the harvest intensity with respect to *Acacia mangium* because this was the dominant species planted by the forest owners. A linear mixed-effects model was used to describe the afforestation and harvest intensity of the forest owners. Three sets of candidate predictors were used to develop the final multiple predictor models. The best subset selection procedure was used to arrive final optimal models describing afforestation and harvest intensity. In this section of the chapter, the optimal models of this study are compared with the ones in the first study.

With respect to the tree planting model, the first study focused on the general tree planting decisions that include reforestation and afforestation decisions. The second study, in contrast, solely focused on the afforestation decisions. However, both studies share a number of key patterns of interaction. These patterns were grouped into four categories, namely forestland management objectives, labour distribution, institutional factors and market conditions.

With respect to the forestland management objectives, both studies indicated that the reason owning forestland for generating forestry income and total forestland plots were positively correlated with tree planting decisions. The results are logical because the forest owners who were motivated by generating income from forestry related activities own more forestland plots, and planting trees is the most direct way to achieve their objectives.

In the first study the reason owning forestland for creating natural landscapes predictor was positively correlated with the decision to plant trees. But this predictor was not correlated in the second study. In the first study, the forest owners were the ones who reforested their forestland. It means that they already spent many years as forest owners. Therefore, it can be assumed that they had positive experiences with forest-related activities. Additionally, research conducted by Meijer et al. (2015) and Malawi and Duesberg et al. (2014) in Ireland show that the owners who had planted trees in the past had more positive attitudes toward tree planting compared to respondents who had not planted trees on their land. Hence, the positive correlation of the reason owning forestland for creating natural landscapes predictor with tree planting decisions in the first study seems reasonable.

With respect to labour distribution, both studies indicated that the age of the forest owners was negatively correlated with decisions in planting trees. It is logically understandable because tree planting activities are physically labour-intensive activities. Therefore, age is a physical barrier preventing the older people from participating in planting trees.

In the first study, the average annual income was negatively correlated with tree planting. Conversely, it was positive correlated with afforestation intensity in the second study. A possible explanation for this might be that afforestation indicates the expansion of new forests and forestry investment requires long-term planning. Therefore, the investment for establishing new forests not only requires instant upfront payment for site preparation but also financial backup planning due to risks that can affect their expected returns, such as bushfire or decreases in timber price. Because of these factors, a higher level of income of the forest owners reflect a greater confidence in new investment.

With respect to institutional factors, the models of both studies indicate that the awareness of the government subsidy grants for establishing forests and technical support from forestry extension workers are positively correlated with planting decisions. This reflects the fundamental economic principle that "people respond to incentives" (Mankiw 2018).

With respect to market conditions, the forest owners in the first study included timber price in their decision. In contrast, forest owners in the second study included the cost of buying seedlings in their decisions.

With respect to harvesting decisions, the first study attempted to model the general harvesting decisions. Meanwhile the second study solely focused on the harvest intensity with respect to *Acacia mangium*. However, timber price and tree age were correlated with the harvest decisions in both models.

CHAPTER 6: CONCLUSIONS

In the last few decades, the forest management system in Vietnam switched from a centrally-planned economy to market-oriented economy. Under the current management model, private ownership is recognised and forests and forestland have been allocated to the local population, which now is one of the largest forest-owning groups in the country. Along with this management model transformation, the forest cover of the country also shifted from net deforestation to net reforestation.

In recent years, forestry policies have also tried to balance economic development with environmental protection and conservation. Various advanced economic instruments are being implemented and tested, such as payment for forest environmental services and REDD+. Because of this, current forest management narratives in Vietnam tend to focus on institutional arrangements and benefit distribution schemes.

The central argument for developing this research is that the most important aspect of forest management is the understanding of forest owners' behaviour. This is because the communication between managers and their policy-targeted audiences is a two-way relationship. The managers use policy as an instrument to guide or change public behaviour. On the other hand, the targeted actors can respond to or alter the institutions that no longer serve their interests. Therefore, the managers should be equipped with instruments that allow them to predict the behaviour of their targeted audiences with the purpose of designing more effective policies in the future.

The purpose of this research was to better understand the non-industrial private forest owners' decisions in planting and harvesting trees by using a case study in the Thai Nguyen province, Vietnam. This principal objective was divided into five sub-objectives:

1. Identifying factors affecting the tree planting and harvesting decisions of NIPF owners.
2. Quantifying the importance of the factors affecting the forest owners' decisions.
3. Testing different regression approaches in modelling and quantifying the importance of factors affecting the decisions of the forest owners.
4. Modelling the afforestation and harvest intensity of NIPF owners.
5. Developing policy recommendations for forest policy makers in Vietnam.

This thesis is composed of six themed chapters. Chapter two detailed the evolution of the forestry landscape in Vietnam with the purpose of framing the significance of this research, as well as reviewing similar research in this topic. Chapter three discussed the nature of policy, the conceptual framework of the research and presented a brief introduction to the region of study. Chapter four discussed the achievement of research objectives number one, two and three. Finally, chapter five discussed the achievement research objective number four.

This chapter of the thesis, chapter six, is a discussion of the findings of the research and to posit further policy implications in practice. With respect to tree planting, this chapter solely focuses on discussions about the afforestation intensity. The researcher believes that afforestation behaviour plays a more important role than reforestation behaviour because afforestation behaviour is a better indicator of the expansion of new forests as compared to reforestation behaviour, due to reasons discussed in chapter four

The chapter comprises four sections. Section 6.1 discusses the research methods that were used in this research. Section 6.2 discusses the afforestation model and its policy applications. Section 6.3 discusses the harvest model and its policy applications. Section 6.4 presents the limitations of the study and presents suggestions for the direction of future research in this topic.

6.1. Methodology

This study used the AID framework, which was introduced by Ostrom (2005) for designing and guiding the research. This framework is particularly useful in the analysis of collective actions that involve social structure, rules, and various stakeholders. A major advantage of this framework is that it serves as a tool to simplify the complexity of the problem by combining the relevant factors into a single system.

Three methods were used in the research; namely the best subset selection, Ridge regression, and LASSO; to obtain multiple predictor models and to quantify the importance of individual predictors in the model. The most significant benefit of these three methods is their ability to balance between bias and variance in model selection, since additional penalties are added to the original regression estimators in each case. The best subset selection technique, in combination with an automated procedure of model fitting, is an alternative to traditional stepwise selection.

The best subset selection technique was unable to explicitly illustrate when and how predictors are eliminated from the model. However, Ridge and LASSO regression were able to provide these metrics. LASSO was recommended to be used because LASSO can perform both regression and variable selection.

The methodology used in this study is easily generalisable and is therefore applicable to a wide range of regions or countries. The IAD framework can help the policy makers in conceptualising and identifying relevant factors and actors relevant to the problem of interest and to set the focal level of policy intervention.

6.2. Afforestation Intensity Model

As mentioned in Chapter two, the main factors behind the deforestation-to-reforestation transition in Vietnam is a topic of debate. It was believed that the increase in forest cover over the last few decades in Vietnam was contributed primarily by private forest owner afforestation. Therefore, speeding up forestland privatisation was believed as a way to increase forest cover in Vietnam. However, this point of view has been challenged by studies such as Clement and Amezaga (2009) and Nguyen and Masuda (2018), noting that forestland usage of FLA recipients is currently not well-understood.

Current debate on forest management in Vietnam to some extent is similar to what Byron (2001b) described. That is much of the research and literature has focused on single aspects of forestry without considering other socio-economic factors that may have an impact on forest owners. Therefore, forestry should be placed among other national economy sectors with the purpose of understanding the interconnection between stakeholders.

The findings reported here shed new light on our understanding of this challenging issue. The results show that the age of the forest owners is negatively correlated with afforestation. This finding contradicts previous research. The research conducted by Kulindwa (2016) in Tanzania; Sikor and Baggio (2014) and Dinh et al. (2017) in Vietnam found that age was not significant in their models. However, the research subjects of those studies were households and the interviewees were the heads of the house. The decision of the heads of households may be affected by the collective decision-making process involving other family members. This may be a reason for this difference.

The education level of the forest owners was also found to be negatively correlated with afforestation intensity. It suggests that a higher level of education might result in a better chance of finding off-farm jobs that can provide better incomes. This finding is contrary to a previous study by Ruseva et al. (2015) in USA that suggested that the level of education is positively correlated on tree planting decision. This difference can be explained by the fact that the highly educated forest owners in USA plant trees because they tend to value non-timber forest products more highly, placing greater value on the aesthetic, conservational, and recreational value of forests. Meanwhile, people with high level of education in Vietnam tend to look for an off-farm job in an urbanized area.

The income level of the forest owners was also found to be positively correlated with afforestation decisions. This finding was also reported by Beach et al. (2005) in USA and Sikor and Baggio (2014) in Vietnam.

The study found that the total cropland area of the forest owners is negatively correlated with afforestation decisions. This outcome is contrary to that of Frayer et al. (2014) in China who found that total cropland area resulted in the increase of forested area. This inconsistency may be due to the different target subjects for the research. The research subjects of Frayer et al. (2014) were households.

The results also indicate that the total number of forest plots owned by the forest owners and the total number of people who participate in forestry activities in forest owners' family are positively correlated with afforestation intensity. This results are in line with findings of Ruseva et al. (2015) in USA, Duesberg et al. (2014) in Ireland, Kulindwa (2016) in Tanzania and Tran et al. (2019) in Vietnam.

This study found that institutional support such as technical support from forestry extension workers and subsidies for establishing forests is positively correlated with the afforestation decisions of forest owners, which was an expected result because afforestation requires a significant upfront payment for site preparation, buying seedlings and planting. Thus, the availability of this support creates incentives for the forest owners to plant trees on their forestland. These results are in agreement with those reported by Frayer et al. (2014) in China, Ruseva et al. (2015) in USA, Kulindwa (2016) in Tanzania, Sikor and Baggio (2014), and Dinh et al. (2017) in Vietnam.

It was found that the most important factors affecting the afforestation decisions of the forest owners are the forestland management objectives. The optimal afforestation models indicated that the forest owners who wanted to generate income from forestry related activities tended to plant trees. Meanwhile, the forest owners who considered owning forestland as an investment were less likely to plant trees. This result suggests that owning forestland does not naturally imply that the owners would be willing to plant trees.

This result also suggests that forestland allocation can negatively impact the total size of the forested area. It must be noted that forestland-use rights include access, exclusion, withdrawal, management and alienation. The access, exclusion, withdrawal and management rights create incentives for the forest owners to afforest their forestland because it ensures that they can freely utilise their forestland and enjoy the products that derived from their forestland.

However, the alienation right allows a person to leave or to enter the forestry sector by transferring forestland use rights to or obtaining use rights from others. If forestland-use rights are transferred to investors who only consider forestland as an investment and may sell the usage rights at their discretion, they have less incentive to plant trees. Additionally, as the current generation of forest owners is getting older, they are faced with the pressure of transferring the forestland-use rights to the next generation, whose behaviour might be markedly different from the current one, leading to uncertainties about future trends in forest cover.

From the above discussions, it can be concluded that allocating forestland to individuals is an appropriate approach to create incentives for tree planting. Additionally, the use of technical and financial support for tree planting activities is necessary to increase the total forest cover. Recent research on forest management in Vietnam, includes discussions on institutional arrangements (McElwee 2012; Nguyen et al. 2018; Suhardiman et al. 2013), benefit distribution in PFES schemes (Duong and de Groot 2018; Hoang et al. 2013; Pham et al. 2015) and social safeguards (Nielsen et al. 2018).

Discussions on payment for forest environmental service are currently focusing protection and special-use forest. No evidence in the country indicates that this payment is constructed for production forests most of which area is planted with acacia, a short rotation species designed for wood products.

However, there is a movement in law and policies that may create new incentives for forest owners planting trees. Firstly, the new Forestry Law introduced two new terms (i) leasing forest environment and (ii) payment for environmental service that include paying for carbon off-set. Secondly, the government is encouraging forest owners to convert small diameter tree production with short rotations to the production of large diameter trees with long rotations. This may imply that in future the forest environmental values will be added into the market.

The efforts of the government and researchers in establishing markets, lifting regulation and providing technologies to improve forest quantity and quality should be recognised and appreciated. However, currently missing from the discussion on the forestry landscape in Vietnam is the understanding of the forestland management objectives of the forest owners. The decisions of the forest owners are shaped by various factors and reflect economic principles. Therefore, the forest policy makers should be equipped with instruments that can help them in observing and identifying the pattern of changes in the forestry landscape to design policies that can influence these changes.

To achieve this goal, it is necessary to develop a management instrument that allows the forest policy makers to (i) understand their policy targeted audiences, (ii) test the impact of their policy during the policy designing stage, and (iii) evaluating impacts of policies by observing changes in society, such as the overall labour distribution, the structure of the population and market conditions. The researcher believes that the approach that was developed in this comprehensive study is applicable for this purpose.

Several policy implications can be drawn from this empirical study:

1. Increasing the rate of forestland allocation to individuals.
2. Continuing to apply tree-planting subsidy schemes with the purpose of creating financial incentives for tree planting.
3. Observing changes in the demographic structure, including the age, education level and income level of the forest owners, as well as the labour distribution in the forestry sector. These factors are indicators of the structure of the income of the forest owners.

6.3. Harvest Intensity Model

With respect to the harvest intensity, the optimal model indicates that timber price is positively correlated with harvesting decisions. This finding is in agreement with findings of van Putten and Jennings (2010) in Tasmania, Australia. The impact of deviation of tree age is in agreement with the work of Sein and Mitlöhner (2011) with respect to *Acacia mangium* in Vietnam.

The negative impact of gender on harvest intensity is logical in the context of rural areas in Vietnam. Traditionally, Vietnamese women tend to focus on domestic activities. This is the barrier preventing them from accessing market information.

6.4. Limitations and future research

In the first study it was concluded that LASSO regression was the superior choice for the purposes of this research. However, the second study used best subset selection instead of LASSO to arrive at multiple predictor models. In fact, several solutions were tried to combine LASSO and mixed-effect models in R. Unfortunately, no feasible solution has been achieved. Therefore, a greater focus on this combination in the future could produce interesting findings.

As mentioned earlier, the researcher believes that the afforestation behaviour plays a more important role than reforestation behaviour because afforestation behaviour is a better indicator of the expansion of new forests as compared to reforestation behaviour. However, further work that combines reforestation and afforestation in a single model would be useful in investigating the relationship between these two types of behaviours. One possible method would be to use reforestation and afforestation indicators as a random effect in the mixed effect model.

It was previously mentioned that the most important factor affecting the behaviour of the forest owners is the forestland management objectives. The presence of this factor in the model significantly improve the quality of the model. However, the information about forest management objectives is directly collected from the forest owners. Therefore, it is challenging for managers use their in-house management information to draw inferences about the forestland management objectives of the forest owners. Because of this, future research on modeling the forestland management objectives of the forest owners based on current governmental databases on the forest owners would be useful./.

Appendix 1: List of variables for the first study

Variable groups	Variable ID	Description
Dependent Variable for Planting Model	Tplnt	Tree-planting decision categorical variable
Dependent Variable for Harvest Model	thvt	Tree-harvesting decision categorical variable
Independent Variables		
Demographic Characteristics	Age	Age of forest owners
	Gen	Gender of forest owners
	Edut	Education level of forest owners
	Fco	Total number of people who living together with forest owners
	FcoM	Total number of males living in a house with the forest owners
	FcoF	Total number of females living in a house with the forest owners
	FcoW	Total number people in the workforces living in a house with the forest owner family (above 15-year-old)
	FcoPF	Total number people participating in forestry jobs living in a house with the forest owner's family
Land Asset	Tfp	Total number of forestland plots of the forest owners
	Parcel	Total forest and forestland area of the forest owners
	crpha	Total cropland area of the forest owners
Annual Income Structure	ConLin	The average annual income of the forest owners
	Crpin	The importance of crop income to annual income of the forest owners
	Frtin	The importance of forestry income to annual income of the forest owners
	Othin	The importance of off-farm income to annual income of the forest owners
Forest Land Management Objectives	Ronlc	The reason for owning forestland to create natural landscape
	Rogfi	The reason owning forestland to generate forestry income
	Roinv	The reason for owning forestland as an investment
	Rokfg	The reason for owning forestland as a way keeping land for future generations
	Fmain	Frequency of forest maintenance activities
For Tree Planting Model		
Market Conditions for Tree Planting	Bseed	Impact of seedling cost on the tree-planting decisions
	Seed	Impact of the availability of seedling supply on the tree-planting decisions
	Hppl	Impact of the cost of hiring people for planting trees on the tree-planting decisions
	TprP	Impact of timber price on the tree-planting decisions
Physical conditions	Scond	Impact of soil conditions on the tree-planting decisions
	Psize	Impact of forestland plot size on the tree-planting decisions
Family labour distribution	FmavP	Impact of availability of family member on tree-planting decisions
Institutional Supports	Ugovt	The use of the government grant for tree planting
	Govt	Impact of awareness of government grants on tree planting decisions
For Harvesting Model		
Tree Biological Characteristics	Tage	The age of trees
Personal Preferences	Csave	Keeping trees as type of capital savings
Institutional Regulation	Greg	Regulations on tree harvesting activities
Market Conditions for Tree Harvesting	Chvt	Cost of harvesting trees
	TprH	Timber price affecting forest owner tree-harvesting decisions

Appendix 2: Function selection

The final goal of constructing the statistical predictive models is to find a useful function representing the relationship between the responses and predictors (Hastie et al. 2009). A common choice for a starting function is a linear function. Fitting a linear function produces a straight line for a set of data (Gravetter and Wallnau 2009) is called multiple predictor linear regression:

Equation 5: Multiple predictor Linear Regression

$$Y = \beta_0 + \sum_{j=1}^{p} \beta_j X_j + \varepsilon$$

where:

- Y is a variable of primary interest, the outputs
- β_0, β_j are unknown parameters that are needed to be estimate from the data
- ε is the residuals that represents factors other than X_j affecting Y.
- X_j j = (1,…, p) is an input vector of p predictors

Equation 5 indicates that the response, Y, can vary on a scale of $(-\infty, +\infty)$ as X_j can vary on a scale of $(-\infty, +\infty)$. However, as defined in Table 2, the value of Y in this study can only vary on a scale of (0, 1). Thus, the Equation 5 has to satisfy the two critical conditions given in Equation 6

Equation 6: The Order Conditions

$$\begin{cases} \left(\beta_0 + \sum_{j=1}^{p} \beta_j X_j + \varepsilon\right) \geq 0 & (3.1) \\ \left(\beta_0 + \sum_{j=1}^{p} \beta_j X_j + \varepsilon\right) \leq 1 & (3.2) \end{cases}$$

The condition (3.1) and (3.2) will collectively be met by Equation 7

Equation 7: Conditional Mean of Y Given X

$$0 \leq \pi(x) = \frac{e^{\left(\beta_0 + \sum_{j=1}^{p} \beta_j X_j + \varepsilon\right)}}{1 + e^{\left(\beta_0 + \sum_{j=1}^{p} \beta_j X_j + \varepsilon\right)}} \leq 1$$

Substituting Equation 7 into Equation 5 , we have:

$$\mathrm{Y} = \frac{e^{\left(\beta_0+\sum_{j=1}^{p}\beta_j X_j+\varepsilon\right)}}{1+e^{\left(\beta_0+\sum_{j=1}^{p}\beta_j X_j+\varepsilon\right)}}$$

$$\Leftrightarrow \mathrm{Y} * (1 + e^{\left(\beta_0+\sum_{j=1}^{p}\beta_j X_j+\varepsilon\right)}) = e^{\left(\beta_0+\sum_{j=1}^{p}\beta_j X_j+\varepsilon\right)}$$

$$\Leftrightarrow Y = (1 - Y) * e^{\left(\beta_0+\sum_{j=1}^{p}\beta_j X_j+\varepsilon\right)}$$

$$\Leftrightarrow \frac{Y}{(1-Y)} = e^{\left(\beta_0+\sum_{j=1}^{p}\beta_j X_j+\varepsilon\right)}$$

Taking the natural logarithm of both sides, we arrive at:

Equation 8: Multiple predictor Logistic Regression

$$ln\left(\frac{Y}{(1-Y)}\right) = \beta_0 + \sum_{j=1}^{p} \beta_j X_j + \varepsilon$$

Equation 8 is called the multiple logistic regression (Fahrmeir et al. 2013; Harrell 2015; Hastie et al. 2009; Lemeshow et al. 2013). The left-hand side of Equation 8 is called the log-*odd* or *logit*. As can be seen the *logit* is linear in its parameters and varies on a scale of $(-\infty$, $+\infty)$ as X_j varies on a scale of $(-\infty, +\infty)$. The second important feature of a logistic regression is that the conditional distribution of Y follows the binomial distribution with the probability given by the conditional mean $\pi(x)$ (Lemeshow et al. 2013).

In Equation 8, β_0 *and* β_j are unknown parameters, or coefficients, which are estimated based on the data using *maximum likelihood estimator*. After estimating the coefficients, assessment of the statistical significance of the variables in the models is required. The purpose of the assessment is to address a central question *Do the variables remaining in the model have an impact on the value of* Y? The classical method for testing statistical significance of the variable is to use *Hypothesis testing.* Two *Hypotheses* will be stated as below:

1. *Null Hypothesis, H_0: The variable has no impact on the response*
2. *Alternative Hypothesis, H_1: The variable has an impact on the response*

The decision on whether to reject or accept the H_0 will be based on the *p-value.* A common practice is the rejection of H_0 and the accepting of H_1 if the p-value is less than 0.05, and accepting of H_0 and the rejection of H_1 if the p-value is greater than 0.05. However, at this stage of the research the goal is to identify candidate predictors for the final statistical predictive models. Hence, to increase the chance of a predictor being included in the model, the threshold for *p-value* was set at 0.1 instead of 0.05.

The mathematical details of *maximum likelihood estimator* and *assessing statistical significance* of the variables are beyond the scope of this thesis. These details can be found in Fahrmeir et al. (2013), Gravetter and Wallnau (2009), Harrell (2015), Hastie et al. (2009), James et al. (2013), Lemeshow et al. (2013), and Gelman and Hill (2007).

Appendix 3: The best subset selection and the use of the AIC

The multiple-predictive models are created by considering all possible combinations from the set of *k* candidate predictors. Mathematically, if there are *k* candidate predictors, there will be 2^k possible combinations for forming models. Classically, backward or forward stepwise selections are often the most ubiquitous choices for generating a final predictive model. However, these methods have disadvantages. Forward stepwise selection does not ensure identification of the best model out of 2^k possible combinations. Meanwhile, backward stepwise selection requires that the number of observations be larger than number of predictors. Details about forward and backward stepwise selection can be found in James et al. (2013), and Miller (2002).

In the regression setting, the Mean Square Error (MSE) is one of the chosen criteria for determining the best-fit models. Meanwhile, in a logistic regression, the fraction of misclassifications criterion is employed. Although these criteria have different names, they are equivalent in nature. These criteria represent the differences between the prediction value and the real value of the response. In other words, these criteria can used to measure the accuracy of the statistical models. To calculate these criteria, the cross-validation set approach and in-sample prediction error are often used.

The cross-validation set approach involves dividing a dataset of n observations into two sets: the validation set of k observations and a training set of n-k observations. If k = 1, the approach is called Leave-One-Out-Cross-Validation (LOOCV). If k>1, the approach is called k-Fold Cross Validation. The cross-validation set approach is more suitable in a data-rich environment, for example in the data mining industry.

However, according to Harrell (2015), the split may cause random errors due to the unpredictable nature of the splitting. If the splitting process is repeated many times, a different predictive accuracy of the model could be obtained. Because of this limitation, this research uses the in-sample prediction error method for assessing the statistical accuracy of the final model.

According to Hastie et al. (2009), James et al. (2013), and Miller (2002), the in-sample method uses the sample to fit or train the model. The statistical accuracy is estimated by making an adjustment to the training error, accounting for bias due to overfitting. The criteria that can be used include Mallow’s Cp, Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC) and adjusted R^2.

In this study, the AIC criterion was chosen for measuring the statistical accuracy of the final model because of two reasons. Firstly, AIC is one of the key model outputs by default in most statistical software. Secondly, according to Stone (1977) and Golub et. Al (1979), as cited in Burham and Anderson (2002), there is an asymptotic equivalence of the choice of model by cross-validation and Akaike's criterion (AIC) when a maximum likelihood estimator is employed within each model.

Appendix 4: Information theoretical approach

The details about AIC can be found in Burnham and Anderson (2002), Burnham et al. (2011), Hastie et al. (2009), James et al. (2013) and Stone (1977). This appendix is a brief synthesis of the Akaike Information Criterion (AIC) from the above sources.

What is AIC?

For a logistic regression model, using a binomial log-likelihood, the mathematical equation of the AIC is:

Equation 9: AIC equation (Adapted from Hastie et al., 2009, p.231)

$$\text{AIC} = -\frac{2}{N}*\text{loglik} + 2*\frac{d}{N}$$

Where:

N is number of observations and

d is number of fitted parameters in the model, including the intercept.

What is the AIC difference (Δ_i)?

For a set of R candidate models, the AIC difference *(Δ_i)* is defined as the difference between the AIC of model i and minimum AIC value in the set.

Equation 10: AIC difference

$$\Delta_i = AIC_i - AIC_{min}$$

What is Akaike Weights (w_i)?

Akaike weights (also called Evidence weights), w_i, takes a value between 0 and 1, with the sum of all Akaike weights of all candidate models in the set being 1. It is interpreted as the probability of a model being the best model in a set of R candidate models.

For example, the Akaike weight of a model in the set of candidate models is 0.3. This is equivalent to saying that there is 30% chance that the model is the best approximating model that describes the data, given the set of candidate models that is being considered.

The Akaike weight is calculated as:

Equation 11: Akaike weights

$$w_i = \frac{exp\left(\frac{-1}{2}i\right)}{\sum_{r=1}^{R} exp\left(\frac{-1}{2}i\right)}$$

What is Multi-model inference?

Multi-model inference produces parameter and error estimates that are not conditional on any model but instead derive from the weighted averages of these values across multiple models. The averaged model's parameters are calculated as:

Equation 12: Averaged model's parameter estimation

$$\hat{\beta_\imath} = \frac{\sum_{r=1}^{R} \beta_i w_i}{\sum_{r=1}^{R} w_i}$$

Where: β_i: Estimate for a predictor in a given model i in the set of R candidate models.

$\beta_i = 0$ if the predictor does not appear in the model.

w_i: Akaike weight of that model in the set.

Appendix 5: List of variables for the second study

Variable group	Variable	Description
Dependent Variable for Afforestation Intensity Model	yearpro	Afforestation intensity, proportion of afforested area versus total land available for afforestation in the year
Dependent Variable for Harvest Intensity Model	hvpro	Harvest intensity, proportion of harvested area versus total available forest area for harvesting in the year
Independent Variables for Afforestation and Harvesting Model		
Demographic Characteristics	age	Age of the forest owners
	ybfo	Number of years being forest owners
	gen	Gender of the forest owners
	edut	Education level of the forest owners (categorical, = 1 if having high school or higher education degree)
	fco	Total number of people who living together with the forest owners
	fcom	Total number of males in the forest owner's family
	fcof	Total number of females in the forest owner's family
	fcow	Total number people who are in workforce in the forest owner's family (above 15 year old)
	fcopf	Total number people who participating in forestry jobs in the forest owner's family
Land Asset	tfp	Total number of forestland plot of the forest owners
	tfl	Total forestland area of the forest owners
	crpha	Total crop land area of the forest owner
Annual Income Structure	llin	The natural logarithm of the forest owner annual income
	crpin	Percentage contribution of crop income to annual income
	frtin	Percentage contribution of forestry income to annual income
	othin	Percentage contribution of off-farm income to annual income
Forest Land Management Objectives	ronlc	Reason owning forestland to create natural landscape
	rogfi	Reason owning forestland to generate forestry income
	roinv	Reason owning forestland as an investment. Land may be sold in future if someone offers the forest owners a reasonable price.
For Afforestation Model		
Institutional Supports	grntaw	The awareness of forest owner about government tree planting grant (categorical, 0: did not know about the grant; 1: know about the grant)
	sup	Receiving technical supports from forestry extension workers

Variable group	**Variable**	**Description**
		(categorical, 0: did not receive the technical support; 1: did receive the technical support)
Market Conditions for Tree Planting		
	tpr	Timber price affects the forest owner tree planting decision (categorical, 0: No; 1: Yes)
	rtpr	The actual market price of timber
	seed	Cost of buying seedlings affects tree planting decision (categorical, 0: No; 1: Yes)
	rseed	The market price of seedlings
	fert	Cost of buying fertilizer affects tree planting decision (categorical, 0: No; 1: Yes)
	rfert	Market price of fertilizer at the time
	cplnt	Cost of planting trees affects tree planting decision (categorical, 0: No; 1: Yes)
For Harvesting Model		
Tree Biological Characteristics	rtage	The age of trees that forest owner harvest
Personal Preferences	fsave	Keeping trees as type of capital save
Institutional Regulation	govhar	Regulation of
Market Conditions for Tree Harvesting	char	Cost of harvesting trees
	tprhar	Timber price affects forest owner tree harvesting decision

REFFERENCES

Bartlett, A.G. 2016. The evolution and impacts of ACIAR’s forestry research program over three decades. Aust Forest 79(3):171-188.

Beach, R.H., S.K. Pattanayak, J.-C. Yang, B.C. Murray, and R.C. Abt. 2005. Econometric studies of non-industrial private forest management: a review and synthesis. Forest Policy and Economics 7(3):261-281.

Boulay, A., L. Tacconi, and P. Kanowski. 2012. Drivers of adoption of eucalypt tree farming by smallholders in Thailand. Agroforestry Systems 84(2):179-189.

Byron, N. 2001a. KEYS TO SMALLHOLDER FORESTRY. Forests, Trees and Livelihoods 11(4):279-294.

Byron, N. 2001b. Keys to smallholder forestry. For. Trees Livelihoods 11(4):279-294.

Castella, J.C., S. Boissau, N.H. Thanh, and P. Novosad. 2006. Impact of forestland allocation on land use in a mountainous province of Vietnam. Land Use Policy 23(2):147-160.

Clement, F., and J.M. Amezaga. 2009. Afforestation and forestry land allocation in northern Vietnam: Analysing the gap between policy intentions and outcomes. Land Use Policy 26(2):458-470.

Clement, F., D. Orange, M. Williams, C. Mulley, and M. Epprecht. 2009. Drivers of afforestation in Northern Vietnam: Assessing local variations using geographically weighted regression. Applied Geography 29(4):561-576.

Colebatch, H.K. 2002. Policy. Open University Press, Philadelphia, PA.

Conway, M.C., G.S. Amacher, J. Sullivan, and D. Wear. 2003. Decisions nonindustrial forest landowners make: an empirical examination. Journal of Forest Economics 9(3):181-203.

Côté, M.-A., A. Généreux-Tremblay, D. Gilbert, and N. Gélinas. 2017. Comparing the profiles, objectives and behaviours of new and longstanding non-industrial private forest owners in Quebec, Canada. Forest Policy and Economics 78(Supplement C):116-121.

Dang, T.K.P., E. Turnhout, and B. Arts. 2012. Changing forestry discourses in Vietnam in the past 20years. Forest Policy and Economics 25:31-41.

Dang, T.K.P., I.J. Visseren-Hamakers, and B. Arts. 2018. Forest devolution in Vietnam: From rhetoric to performance. Land Use Policy 77:760-774.

Dao Minh, T., M. Yanagisawa, and Y. Kono. 2017. Forest transition in Vietnam: A case study of Northern mountain region. Forest Policy and Economics 76:72-80.

Dennis, D.F. 1989. An economic analysis of harvest behavior: integrating forest and ownership characteristics. Forest Science 35(4):1088-1104.

Dermawan, A., E. Kemp-Benedict, A. Huber-Lee, and A. Fencl. 2013. Testing a multi-scale scenario approach for smallholder tree plantations in Indonesia and Vietnam. Technological Forecasting and Social Change 80(4):762-771.

Dinh, H.H., T.T. Nguyen, V.N. Hoang, and C. Wilson. 2017. Economic incentive and factors affecting tree planting of rural households: Evidence from the Central Highlands of Vietnam. Journal of Forest Economics 29:14-24.

Do, T.H., T.P. Vu, V.T. Nguyen, and D. Catacutan. 2018. Payment for forest environmental services in Vietnam: An analysis of buyers' perspectives and willingness. Ecosystem Services 32:134-143.

Douglas Bates, M.M., Ben Bolker, Steve Walker 2015. Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software, 67(1):1-48.

Duesberg, S., V. Upton, D. O'Connor, and T.N. Dhubháin. 2014. Factors influencing Irish farmers' afforestation intention. Forest Policy and Economics 39:13-20.

Duong, N.T.B., and W.T. de Groot. 2018. Distributional risk in PES: Exploring the concept in the Payment for Environmental Forest Services program, Vietnam. Forest Policy and Economics 92:22-32.

Ficko, A., and A. Boncina. 2013. Probabilistic typology of management decision making in private forest properties. Forest Policy and Economics 27:34-43.

Flanagan, A.C., S.J. Midgley, P.R. Stevens, and L. McWhirter. 2019. Smallholder tree-farmers and forest certification in Southeast Asia: productivity, risks and policies. Aust Forest 82(1):18-28.

Fox, J., and S. Weisberg. 2011. An {R} Companion to Applied Regression, Second Edition. Sage, Thousand Oaks CA.

Frayer, J., Z. Sun, D. Müller, D.K. Munroe, and J. Xu. 2014. Analyzing the drivers of tree planting in Yunnan, China, with Bayesian networks. Land Use Policy 36(0):248-258.

Frey, G.E., F.W. Cubbage, T.T.T. Ha, R.R. Davis, J.B. Carle, V.X. Thon, and N.V. Dzung. 2018. Financial analysis and comparison of smallholder forest and state forest enterprise plantations in Central Vietnam. Int. For. Rev. 20(2):181-198.

Gelman, A., and J. Hill. 2007. Data analysis using regression and multilevel/hierarchical models. Cambridge University Press, Cambridge;New York;.

Gravetter, F.J., and L.B. Wallnau. 2009. Statistics for the behavioral sciences. Wadsworth, Belmont, CA.

Gregory, S.A., M. Christine Conway, J. Sullivan, and S.A. Gregory. 2003. Econometric analyses of nonindustrial forest landowners: Is there anything left to study? Journal of Forest Economics 9(2):137-164.

Harrell, F.E. 2015. Regression modeling strategies: with applications to linear models, logistic and ordinal regression, and survival analysis. Springer, Cham.

Harrison, X.A., L. Donaldson, M.E. Correa-Cano, J. Evans, D.N. Fisher, C.E.D. Goodwin, B.S. Robinson, D.J. Hodgson, and R. Inger. 2018. A brief introduction to mixed effects modelling and multi-model inference in ecology. PeerJ 6:e4794.

Harwood, C.E., E.B. Hardiyanto, and W.C. Yong. 2015. Genetic improvement of tropical acacias: achievements and challenges. Southern Forests: a Journal of Forest Science 77(1):11-18.

Harwood, C.E., E.K.S. Nambiar, P.X. Dinh, L.X. Toan, and L.T. Quang. 2017. Managing wood production from small grower acacia hybrid plantations on eroded soils in central Vietnam. Aust Forest 80(5):286-293.

Harwood, C.E., and N.K.S. Nambiar. 2014. Sustainable plantation forestry in South-East Asia.

Hengeveld, G.M., E. Schüll, R. Trubins, and O. Sallnäs. 2017. Forest Landscape Development Scenarios (FoLDS)–A framework for integrating forest models, owners' behaviour and socio-economic developments. Forest Policy and Economics.

Hlavac, M. 2018. stargazer: Well-Formatted Regression and Summary Statistics Tables. R package version 5.2.1. https://cran.r-project.org/package=stargazer.

Hoang, M.H., T.H. Do, M.T. Pham, M. van Noordwijk, and P.A. Minang. 2013. Benefit distribution across scales to reduce emissions from deforestation and forest degradation (REDD+) in Vietnam. Land Use Policy 31:48-60.

Howard, C. 1998. Forestry in transition in Vietnam. Commonw. For. Rev.(251):249-253.

Hox, J.J., and J.K. Roberts. 2011. Handbook of Advanced Multilevel Analysis. Routledge.

Irimie, D.L., and H.F. Essmann. 2009. Forest property rights in the frame of public policies and societal change. Forest Policy and Economics 11(2):95-101.

Jakobsen, J., K. Rasmussen, S. Leisz, R. Folving, and N.V. Quang. 2007. The effects of land tenure policy on rural livelihoods and food sufficiency in the upland village of Que, North Central Vietnam. Agricultural Systems 94(2):309-319.

Joshi, S., and K.G. Arano. 2009. Determinants of private forest management decisions: A study on West Virginia NIPF landowners. Forest Policy and Economics 11(2):118-125.

Khuc, Q.V., B.Q. Tran, P. Meyfroidt, and M.W. Paschke. 2018. Drivers of deforestation and forest degradation in Vietnam: An exploratory analysis at the national level. Forest Policy and Economics 90:128-141.

Kooreman, P., and S. Wunderink. 1997. The economics of household behaviour. Macmillan, Houndmills, Basingstoke.

Kruschke, J.K. 2011. Doing Bayesian data analysis: a tutorial with R and BUGS. Academic Press, Amsterdam;Boston;.

Kulindwa, Y.J. 2016. Key factors that influence households' tree planting behaviour. Natural Resources Forum 40(1-2):37-50.

Lambini, C.K., T.T. Nguyen, J. Abildtrup, V.D. Pham, J. Tenhunen, and S. Garcia. 2018. Are Ecosystem Services Complementary or Competitive? An Econometric Analysis of Cost Functions of Private Forests in Vietnam. Ecological Economics 147:343-352.

Lavrakas, P.J. 2008. Encyclopedia of Survey Research Methods.

Lindley, D.V. 1971. Making decisions. Wiley-Interscience, New York; London.

Lüdecke, D. 2018. sjPlot: Data Visualization for Statistics in Social Science. https://doi.org/10.18637/jss.v067.i01.

Mankiw, N.G. 2018. Principles of economics. Cengage Learning.

McElwee, P. 2004. You say illegal, I say legal: The relationship between 'illegal' logging and land tenure, poverty, and forest use rights in Vietnam. J. Sustainable For. 19(1-3):97-135.

McElwee, P.D. 2012. Payments for environmental services as neoliberal market-based forest conservation in Vietnam: Panacea or problem? Geoforum 43(3):412-426.

Meijer, S.S., D. Catacutan, G.W. Sileshi, and M. Nieuwenhuis. 2015. Tree planting by smallholder farmers in Malawi: Using the theory of planned behaviour to examine the relationship between attitudes and behaviour. Journal of Environmental Psychology 43:1-12.

Meyfroidt, P., and E.F. Lambin. 2008. The causes of the reforestation in Vietnam. Land Use Policy 25(2):182-197.

Midgley, S.J., P.R. Stevens, and R.J. Arnold. 2017a. Hidden assets: Asia’s smallholder wood resources and their contribution to supply chains of commercial wood. Aust Forest 80(1):10-25.

Midgley, S.J., P.R. Stevens, R.J. Arnold, and R.J. Arnold. 2017b. Hidden assets: Asia's smallholder wood resources and their contribution to supply chains of commercial wood. Aust Forest 80(1):10-25.

Ministry Of Agriculture and Rural Development. 2017. QD 1819/QD-BNN-TCLN dated on 16/5/2017 Annoucement of Forest Status in 2016. vietnam.

Musshoff, O., and S.C. Maart-Noelck. 2014. An experimental analysis of the behavior of forestry decision-makers — The example of timing in sales decisions. Forest Policy and Economics 41(0):31-39.

Naimi, B., N.A.S. Hamm, T.A. Groen, A.K. Skidmore, and A.G. Toxopeus. 2014. Where is positional uncertainty a problem for species distribution modelling? Ecography 37(2):191-203.

Nguyen, M.D., T. Ancev, and A. Randall. 2018. Forest governance and economic values of forest ecosystem services in Vietnam. Land Use Policy.

Nguyen, T.Q. 2006. Forest devolution in Vietnam: Differentiation in benefits from forest among local households. Forest Policy and Economics 8(4):409-420.

Nguyen, T.T., S. Bauer, and H. Uibrig. 2010. Land privatization and afforestation incentive of rural farms in the Northern Uplands of Vietnam. Forest Policy and Economics 12(7):518-526.

Nguyen, T.T., and M. Masuda. 2018. Land Use After Forestland Allocation and the Potential for Farm Forestry in a Mountainous Region of Northeast Vietnam. Small-scale For.:1-19.

Ní Dhubháin, Á., K. Maguire, and N. Farrelly. 2010. The harvesting behaviour of Irish private forest owners. Forest Policy and Economics 12(7):513-517.

Nielsen, M.R., I. Theilade, H. Meilby, N.H. Nui, and N.T. Lam. 2018. Can PES and REDD+ match Willingness To Accept payments in contracts for reforestation and avoided forest degradation? The case of farmers in upland Bac Kan, Vietnam. Land Use Policy 79:822-833.

Novais, A., and M.J. Canadas. 2010. Understanding the management logic of private forest owners: A new approach. Forest Policy and Economics 12(3):173-180.

Ostrom, E. 2005. Understanding institutional diversity. Princeton University Press, Princeton, NJ.

Pham, T.T., L. Loft, K. Bennett, V.T. Phuong, L.N. Dung, and J. Brunner. 2015. Monitoring and evaluation of Payment for Forest Environmental Services in Vietnam: From myth to reality. Ecosyst. Serv. 16:220-229.

R Core Team. 2017. R: A language and environment for statistical computing. in https://www.R-project.org/, Computing, R.F.f.S. (ed.), Vienna, Austria.

RStudio Team. 2015. RStudio: Integrated Development for R. RStudio, Inc., Boston, MA http://www.rstudio.com/.

Ruseva, T.B., T.P. Evans, and B.C. Fischer. 2015. Can incentives make a difference? Assessing the effects of policy tools for encouraging tree-planting on private lands. Journal of Environmental Management 155(0):162-170.

Sein, C.C., and R. Mitlöhner. 2011. Acacia mangium Willd: ecology and silviculture in Vietnam. CIFOR, Bogor, Indonesia.

Sikor, T., and J.A. Baggio. 2014. Can Smallholders Engage in Tree Plantations? An Entitlements Analysis from Vietnam. World Development 64, Supplement 1:S101-S112.

Sikor, T., and P.X. To. 2011. Illegal logging in vietnam: Lam Tac (Forest Hijackers) in Practice and Talk. Soc. Nat. Res. 24(7):688-701.

Størdal, S., G. Lien, and S. Baardsen. 2008. Analyzing determinants of forest owners' decision-making using a sample selection framework. Journal of Forest Economics 14(3):159-176.

Suhardiman, D., D. Wichelns, G. Lestrelin, and C.T. Hoanh. 2013. Payments for ecosystem services in Vietnam: market-based incentives or state control of resources? Ecosyst. Serv. 6:64-71.

Tachibana, T., T.M. Nguyen, and K. Otsuka. 2001. Agricultural intensification versus extensification: A case study of deforestation in the Northern-hill region of Vietnam. J. Environ. Econ. Manage. 41(1):44-69.

Trædal, L.T., P.O. Vedeld, and J.G. Pétursson. 2016. Analyzing the transformations of forest PES in Vietnam: Implications for REDD+. Forest Policy and Economics 62:109-117.

Tran, T.M.A., D.W. Ko, C.R. Park, and H.D. Le. 2019. A bayesian network analysis of reforestation decisions by rural mountain communities in Vietnam. Forest Science and Technology 15(2):51-57.

van Putten, I., and S. Jennings. 2010. Modeling forest owner harvesting behaviour and future intentions in Tasmania. Small-scale For. 9(2):175-193.

Vincent Calcagno. 2013. glmulti: Model selection and multimodel inference made easy. R package version 1.0.7. in https://CRAN.R-project.org/package=glmulti.

Westing, A.H. 1971. Forestry and the war in South Vietnam. J Forest 69(11):777-783.

Wickham, H. 2017. tidyverse: Easily Install and Load the 'Tidyverse'. R package version 1.2.1.